

# Introduction à la théorie ergodique et aux systèmes dynamiques topologiques et différentiables

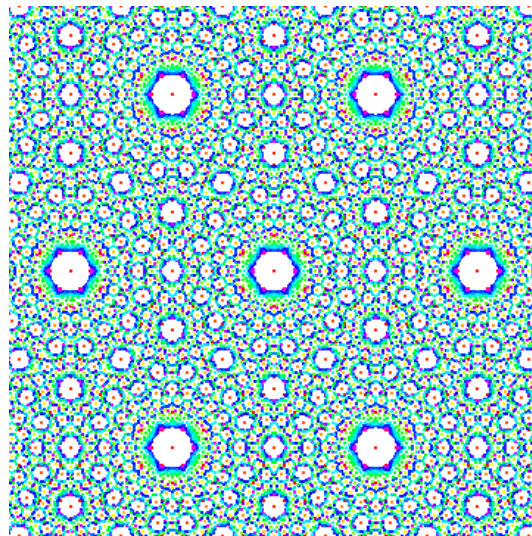
**Frédéric Paulin**

**Professeur à la Faculté des sciences d'Orsay**

Cours de seconde année de Master, parcours :  
Analyse, arithmétique et géométrie (AAG)  
et Mathématiques de l'aléatoire (Mda)

**Université Paris-Saclay**

Année 2023-2024, version préliminaire



# Table des matières

Préambule . . . . .	5
<b>0 Vocabulaire des systèmes dynamiques et exemples fondamentaux</b>	<b>6</b>
0.1 Orbites et orbites périodiques . . . . .	8
0.2 Parties invariantes et mesures invariantes . . . . .	11
Systèmes de Kronecker <sup>27</sup> . . . . .	14
Variables aléatoires indépendantes identiquement distribuées . . . . .	15
Mesures de Liouville <sup>27</sup> et systèmes hamiltoniens . . . . .	17
0.3 Outils d'analyse fonctionnelle . . . . .	20
0.4 Exercices . . . . .	23
0.5 Indications pour la résolution des exercices . . . . .	23
<b>I Théorie ergodique</b>	<b>24</b>
<b>1 Généralités sur la dynamique mesurée</b>	<b>24</b>
1.1 Récurrence mesurée, errance mesurée, transience, conservativité . . . . .	24
1.2 Temps de retour et théorème de Kac <sup>44</sup> . . . . .	25
1.3 Exercices . . . . .	27
1.4 Indications pour la résolution des exercices . . . . .	27
<b>2 Ergodicité</b>	<b>28</b>
2.1 Transformations ergodiques . . . . .	29
2.2 Le théorème ergodique de Birkhoff <sup>50</sup> . . . . .	30
2.3 Le théorème ergodique sous-additif de Kingman <sup>57</sup> . . . . .	35
2.4 Le théorème ergodique multiplicatif d'Oseledets . . . . .	42
2.5 Exercices . . . . .	54
2.6 Indications pour la résolution des exercices . . . . .	59
<b>3 Mélange et vitesses de mélange</b>	<b>66</b>
3.1 Mélange fort . . . . .	66
3.2 Mélange exponentiel . . . . .	69
3.3 Un peu de théorie spectrale de l'opérateur de Koopman . . . . .	70
Transformation de Fourier <sup>92</sup> des mesures positives sur le cercle . . . . .	72
3.4 Exercices . . . . .	76
3.5 Indications pour la résolution des exercices . . . . .	78
<b>4 Dynamique linéaire mesurée des tores</b>	<b>84</b>
4.1 Séries de Fourier sur le tore . . . . .	84
4.2 Unique ergodicité . . . . .	86
4.3 Unique ergodicité des translations sur le tore . . . . .	87
4.4 Equirépartition de suites modulo 1 et unique ergodicité . . . . .	88
4.5 Propriétés de mélange des endomorphismes linéaires du tore . . . . .	91
4.6 Exercices . . . . .	94
4.7 Indications pour la résolution des exercices . . . . .	95

<b>5</b>	<b>Entropie mesurée</b>	<b>98</b>
5.1	Information et entropie d'une partition . . . . .	98
	Partitions jointes . . . . .	99
	Raffinement de partitions . . . . .	99
	Indépendance de partitions . . . . .	100
	Égalité presque partout de partitions . . . . .	100
	Raffinement de partitions . . . . .	101
5.2	Entropie d'un système dynamique mesuré . . . . .	107
5.3	L'inégalité de Ruelle <sup>120</sup> . . . . .	113
5.4	Exercices . . . . .	117
5.5	Indications pour la résolution des exercices . . . . .	119
<b>6</b>	<b>Dynamique symbolique markovienne</b>	<b>122</b>
6.1	Chaînes de Markov <sup>126</sup> sur un alphabet fini . . . . .	122
	Matrices à coefficients positifs ou nuls . . . . .	122
	Matrices de transition . . . . .	124
	Matrices de Markov . . . . .	125
6.2	Théorie ergodique de la transformation de Gauss <sup>126</sup> . . . . .	129
	6.2.1 Fractions continues . . . . .	129
	6.2.2 Ergodicité de la transformation de Gauss . . . . .	132
	6.2.3 Mélange de la transformation de Gauss . . . . .	134
6.3	Exercices . . . . .	135
6.4	Indications pour la résolution des exercices . . . . .	135
<b>II</b>	<b>Systèmes dynamiques topologiques et différentiables</b>	<b>139</b>
<b>7</b>	<b>Généralités de dynamique topologique</b>	<b>139</b>
7.1	Compléments sur les orbites périodiques . . . . .	139
7.2	Récurrence topologique . . . . .	142
7.3	L'ensemble (topologiquement) non errant . . . . .	143
7.4	Transitivité (topologique), mélange topologique et minimalité (topologique) . . . . .	144
7.5	Récurrence multiple . . . . .	149
7.6	Exercices . . . . .	152
7.7	Indications pour la résolution des exercices . . . . .	153
<b>8</b>	<b>Dynamique des homéomorphismes du cercle</b>	<b>159</b>
8.1	Nombre de rotation d'un homéomorphisme du cercle . . . . .	160
8.2	Théorie de Poincaré-Denjoy <sup>144</sup> des homéomorphismes du cercle . . . . .	163
8.3	Problème de $C^r$ -conjugaison et petits diviseurs . . . . .	169
8.4	Exercices . . . . .	170
8.5	Indications pour la résolution des exercices . . . . .	171
<b>9</b>	<b>Dynamique sur les espaces homogènes</b>	<b>174</b>
9.1	Le groupe unimodulaire $SL_N(\mathbb{R})$ . . . . .	174
9.2	Représentations unitaires . . . . .	178
9.3	Ergodicité des translations sur les espaces homogènes . . . . .	180
9.4	Mélange des translations sur les espaces homogènes . . . . .	181

9.5	Flots géodésique et horocyclique sur les surfaces hyperboliques . . . . .	185
9.6	Exercices . . . . .	197
9.7	Indications pour la résolution des exercices . . . . .	200
<b>10</b>	<b>Systèmes dynamiques hyperboliques</b>	<b>207</b>
10.1	Endomorphismes linéaires hyperboliques . . . . .	208
10.2	Le théorème de Grobman-Hartman . . . . .	215
10.3	Stabilité structurelle des automorphismes hyperboliques du tore . . . . .	216
10.4	Ensembles hyperboliques de difféomorphismes . . . . .	220
10.4.1	Fibrés vectoriels normés . . . . .	221
10.4.2	Parties hyperboliques d'une variété munie d'un difféomorphisme . . . . .	224
10.5	Le fer à cheval de Smale . . . . .	231
10.6	Le critère des champs de cônes et la stabilité des ensembles hyperboliques . . . . .	235
10.7	Expansivité et lemme de pistage des systèmes dynamiques hyperboliques . . . . .	239
10.8	Le théorème des variétés stables et instables . . . . .	247
10.8.1	Variétés stables des perturbations d'endomorphismes linéaires . . . . .	247
10.8.2	Variétés stables et instables locales des difféomorphismes . . . . .	251
10.8.3	Variétés stables et instables globales des difféomorphismes . . . . .	255
10.9	La structure de produit local . . . . .	257
10.10	Exercices . . . . .	260
10.11	Appendice : rappels de théorie spectrale . . . . .	262
	Sur la complexification des espaces de Banach réels . . . . .	263
10.12	Indications pour la résolution des exercices . . . . .	267
<b>11</b>	<b>Mesures d'entropie maximale</b>	<b>274</b>
11.1	Entropie topologique . . . . .	274
	Recouvrements ouverts . . . . .	274
	Définition de l'entropie topologique . . . . .	276
	Propriétés élémentaires de l'entropie topologique . . . . .	277
11.2	Principe variationnel . . . . .	279
11.3	Entropie topologique des systèmes dynamiques symboliques . . . . .	284
11.3.1	Sous-décalages de type fini . . . . .	284
11.3.2	Problématique de l'équidistribution . . . . .	287
11.3.3	Mesures d'entropie maximale des sous-décalages de type fini . . . . .	289
11.4	Entropie topologique des automorphismes linéaires du tore . . . . .	293
11.4.1	Partition de Markov du tore . . . . .	294
11.4.2	Codage des automorphismes linéaires hyperboliques du tore . . . . .	298
11.4.3	Mesure d'entropie maximale des automorphismes hyperboliques du tore . . . . .	299
11.5	Exercices . . . . .	302
	<b>Index</b>	<b>304</b>
	<b>Références</b>	<b>310</b>

## Préambule.

En mécanique classique, on étudie l'évolution au cours du temps de certains systèmes physiques comme une toupie, un gaz, une étoile et ses planètes ... Si, au temps initial  $t = 0$ , le système est représenté par un point  $x$  de l'espace des phases  $X$ , alors au temps  $t$ , ce système est représenté par un point  $\phi^t(x)$ .

Lorsque les équations différentielles qui régissent ce mouvement sont indépendantes du temps et lorsque les solutions sont définies en tout temps, l'égalité  $\phi^{t+t'} = \phi^t \circ \phi^{t'}$  est vérifiée pour tous les  $t, t'$  dans  $\mathbb{R}$ . On dit que  $\phi = (\phi^t)_{t \in \mathbb{R}}$  est un groupe à un paramètre de transformations de  $X$ .

Lorsque l'on ne dispose pas de formule explicite pour  $\phi^t$ , on cherche à comprendre le comportement de  $\phi^t(x)$  pour  $t$  grand. Une telle étude qualitative a été initiée par Poincaré.

Pour clarifier ces phénomènes, les mathématiciens sont sortis du cadre des équations différentielles, en ne gardant qu'un espace (avec des structures appropriées)  $X$  et un groupe à un paramètre de transformations (avec des propriétés appropriées de préservations de structures)  $(\phi^t)_{t \in \mathbb{R}}$  [ système dynamique à temps continu ] ou encore, en ne gardant qu'un espace  $X$  et une transformation  $\phi$  (pas toujours inversible) de l'espace  $X$  [ système dynamique à temps discret ], toujours avec des structures appropriées sur  $X$  et des propriétés appropriées de préservations de structures par  $\phi$ . Dans ce dernier cas, on s'intéresse au comportement asymptotique des itérations  $\phi^n$  de  $\phi$  quand l'entier positif  $n$  est grand.

Pour mener l'étude qualitative du comportement de  $\phi^t(x)$  pour  $t$  grand, deux cadres s'avèrent particulièrement bien adaptés : celui de la théorie de la mesure [ théorie ergodique ] et celui de la topologie [ système dynamique topologique, voire système dynamique différentiable quand l'espace des phases est une variété différentielle ]. Dans ce cours, nous étudierons successivement (et indépendamment du point de vue de la validation) ces deux points de vue. Bien plus que de clarifier les idées de ce sujet, ces cadres plus larges et plus naturels ont permis leur application à d'autres domaines des mathématiques (théorie des nombres, théorie des groupes, géométrie différentielle). Nous discuterons de certaines de ces applications.

Une référence de base pour ce cours est le livre encyclopédique [KH], dont nous ne traiterons qu'une toute petite partie. Une partie de ces notes est issue des notes de cours [BenP] rédigées avec Yves Benoist, que je remercie.<sup>1</sup> Tous les portraits de mathématiciens qui apparaissent en note de bas de page de ce cours sont extraits de Wikipedia, sauf ceux d'Anosov, Denjoy, Kingman, Nikodym, Renyi et Stone, venant de <https://mathshistory.st-andrews.ac.uk/Biographies>.

Pour tous les  $p, q \in \mathbb{Z}$ , nous noterons  $\llbracket p, q \rrbracket = [p, q] \cap \mathbb{Z}$  l'intervalle entier d'extrémités  $p, q$ . Si  $A$  et  $B$  sont deux ensembles, nous noterons  $A \setminus B = \{x \in A : x \notin B\}$ . Si  $X$  est un espace topologique, nous noterons  $C^0(X; \mathbb{R})$  ou  $C^0(X; \mathbb{C})$  l'espace vectoriel réel ou complexe des fonctions continues réelles ou complexes sur  $X$ . Lorsque  $X$  est compact, nous le munirons sauf mention explicite du contraire de la norme de la convergence uniforme  $f \mapsto \|f\| = \max_{x \in X} |f(x)|$ , qui en fait un espace de Banach.

**Caveat.** Conformément à l'enseignement gaulois de N. Bourbaki, tous les espaces compacts, localement compacts ou  $\sigma$ -compacts sont, par définition, séparés.

---

1. Je remercie Jean Lécureux et les étudiants de l'année 2023-2024 pour leurs corrections, en particulier Jean-Baptiste Stiegler et Siwei Liang.

## 0 Vocabulaire des systèmes dynamiques et exemples fondamentaux

Ce chapitre 0 est commun aux deux parties de ces notes de cours, la partie I surtout consacrée à la théorie ergodique et la partie II surtout consacrée aux systèmes dynamiques topologiques et différentiables, même si les liens sont nombreux.

Rappelons que si  $(Y, \mathcal{C}, \nu)$  est un espace mesuré, si  $(Y', \mathcal{C}')$  est un espace mesurable, et si  $\psi : Y \rightarrow Y'$  est une application mesurable, la *mesure image*  $\psi_*\nu$  de  $\nu$  par  $\psi$  est la mesure sur  $(Y', \mathcal{C}')$  définie par  $\psi_*\nu(C') = \nu(\psi^{-1}(C'))$  pour tout  $C' \in \mathcal{C}'$ . Notons que si  $(Y'', \mathcal{C}'')$  est un autre espace mesurable et  $\psi' : Y' \rightarrow Y''$  est une autre application mesurable, alors

$$(\psi' \circ \psi)_*\nu = \psi'_* \circ \psi_*\nu .$$

Rappelons que si  $(Y, \mathcal{C}, \nu)$  et  $(Y', \mathcal{C}', \nu')$  sont deux espaces mesurés, nous dirons qu'une application  $\psi : Y \rightarrow Y'$  *préserve la mesure* si  $\psi$  est mesurable et si  $\psi_*\nu = \nu'$  (ou de manière équivalente, si  $\nu(f \circ \psi) = \nu'(f)$  pour toute application  $f : Y' \rightarrow [0, +\infty[$  mesurable, voir par exemple [Coh]). Lorsque  $(Y, \mathcal{C}, \nu) = (Y', \mathcal{C}', \nu')$ , nous dirons aussi que  $\nu$  est *invariante* par  $\psi$ .

Rappelons que si  $X$  est un ensemble, un *groupe à un paramètre* sur  $X$  est une famille  $\phi = (\phi^t : X \rightarrow X)_{t \in \mathbb{R}}$  indexée par  $\mathbb{R}$  d'applications de  $X$  dans  $X$  telle que  $\phi^0 = \text{id}$  et  $\phi^{s+t} = \phi^s \circ \phi^t$  pour tous les  $s, t \in \mathbb{R}$ . En particulier, pour tout  $t \in \mathbb{R}$ , l'application  $\phi^t$  est une bijection de  $X$ , d'inverse  $\phi^{-t}$ .

Un *système dynamique*<sup>2</sup> est un couple  $(\widehat{X}, \phi)$  (on parle alors de *système dynamique à temps discret*) ou  $(\widehat{X}, \phi = (\phi^t)_{t \in \mathbb{R}})$  (on parle alors de *système dynamique à temps continu*), où  $\widehat{X}$ , appelé l'*espace des phases* (et noté souvent par abus juste par son ensemble sous-jacent  $X$ ), est

- un espace mesurable  $(X, \mathcal{B})$  – on parle alors de *système dynamique mesurable* –, ou
- un espace mesuré<sup>3</sup>  $(X, \mathcal{B}, \mu)$  – on parle alors de *système dynamique mesuré*, et de *système dynamique probabilisé* lorsque  $\mu$  est une mesure de probabilité –, ou
- un espace topologique  $(X, \mathcal{O})$  – on parle alors de *système dynamique topologique* –, ou
- une variété différentielle (lisse, de dimension  $d$ )<sup>4</sup>  $(X, \mathcal{A})$  – on parle alors de *système*

2. Nous ne considérerons pas dans ce cours des systèmes dynamiques constitués d'une action d'un groupe ou d'un pseudo-groupe sur un espace, pourtant si intéressants.

3. Sauf mention explicite du contraire, toutes les mesures dans ce texte sont des mesures positives, et toutes les mesures sur un espace topologique sont des mesures boréliennes, éventuellement boréliennes complétées. Nous supposerons en effet souvent et implicitement que la tribu (aussi appelée  $\sigma$ -algèbre)  $\mathcal{B}$  est *complète* pour  $\mu$ , c'est-à-dire que toute partie d'un ensemble de  $\mathcal{B}$  de mesure nulle appartient encore à  $\mathcal{B}$ , ce qui ne pose pas de problème.

4. c'est-à-dire un espace topologique métrisable séparable  $X$  muni d'un atlas de cartes  $\mathcal{A}$  de classe  $C^\infty$  de dimension  $d$  maximal. Nous renvoyons par exemple à [Laf, Pau2] pour les définitions de base. Nous n'aurons pas besoin de plus que

- les sous-variétés lisses  $M$  de  $\mathbb{R}^N$  (munies de l'unique atlas de cartes lisse maximal contenant les *paramétrages locaux*  $(U, \varphi)$  où  $U$  est un ouvert de  $\mathbb{R}^d$  et  $\varphi : U \rightarrow \mathbb{R}^N$  une immersion lisse, d'image un ouvert de  $M$ , et qui est un homéomorphisme sur son image) et

- les variétés quotients  $\Gamma \backslash \widetilde{X}$  d'une variété différentielle lisse  $\widetilde{X}$  de dimension  $d$  par un groupe  $\Gamma$  agissant librement et proprement par difféomorphismes lisses sur  $\widetilde{X}$  (munies de l'unique atlas de cartes lisse maximal contenant les  $(U, \pi \circ \varphi)$  où  $(U, \varphi)$  est une carte locale de  $\widetilde{X}$  telle que  $\varphi(U) \cap \gamma\varphi(U) = \emptyset$  pour tout  $\gamma \in \Gamma - \{\text{id}\}$ , et  $\pi : \widetilde{X} \rightarrow X = \Gamma \backslash \widetilde{X}$  est la projection canonique), comme  $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$ .

*dynamique différentiable*,

et où  $\phi : X \rightarrow X$  est une application appelée la *transformation* du système dynamique à temps discret ou  $\phi = (\phi^t : X \rightarrow X)_{t \in \mathbb{R}}$  est un groupe à un paramètre sur  $X$ , appelé le *flot* du système dynamique à temps continu, qui vérifie respectivement

- $\phi$  est mesurable ou  $\phi^t$  est mesurable pour tout  $t \in \mathbb{R}$ ,
- $\phi$  préserve la mesure ou  $\phi^t$  préserve la mesure pour tout  $t \in \mathbb{R}$ ,
- $\phi$  est continue ou l'application  $(t, x) \mapsto \phi^t(x)$  de  $\mathbb{R} \times X$  dans  $X$  est continue,
- $\phi$  est différentiable<sup>5</sup> de classe  $C^k$  ou l'application  $(t, x) \mapsto \phi^t(x)$  de  $\mathbb{R} \times X$  dans  $X$  est différentiable de classe  $C^k$ , où  $k \in \mathbb{N} - \{0\}$ , et on parle de *système dynamique différentiable de classe  $C^k$*  lorsque nous voulons préciser  $k$ .

Sauf lorsqu'il convient de préciser la tribu, la mesure, la topologie ou l'atlas, nous noterons par abus  $(X, \phi)$  au lieu de  $(\widehat{X}, \phi)$  à partir de maintenant.

**Exemples.** (1) (**Inversibilité**) Un système dynamique à temps discret  $(X, \phi)$  est dit *inversible* si l'application  $\phi$  est bijective, et si  $(X, \phi^{-1})$  est encore un système dynamique du même type (mesurable, mesuré, topologique, différentiable de même classe). Nous appellerons *temps 1* d'un système dynamique à temps continu  $(X, (\phi^t)_{t \in \mathbb{R}})$  le système dynamique à temps discret  $(X, \phi^1)$ , qui est du même type et inversible.

(2) (**Appauvrissement de structure**) Si  $(X, \mathcal{B}, \mu, \phi)$  est un système dynamique mesuré, alors  $(X, \mathcal{B}, \phi)$  est un système dynamique mesurable. Si  $(X, \mathcal{O}, \phi)$  est un système dynamique topologique et si  $\mathcal{B}$  est la tribu des boréliens de  $X$ , alors  $(X, \mathcal{B}, \phi)$  est un système dynamique mesurable. Si  $(X, \mathcal{A}, \phi)$  est un système dynamique différentiable et si  $\mathcal{O}$  est la topologie sur  $X$  définie par l'atlas  $\mathcal{A}$  (c'est-à-dire la topologie sous-jacente de la variété différentielle  $(X, \mathcal{A})$ ), alors  $(X, \mathcal{O}, \phi)$  est un système dynamique topologique. Pour tous les entiers  $k \leq k'$ , si  $(X, \phi)$  est un système dynamique différentiable de classe  $C^{k'}$ , alors  $(X, \phi)$  est aussi un système dynamique différentiable de classe  $C^k$ , en munissant  $X$  de l'atlas de cartes maximal de classe  $C^k$  contenant son atlas  $C^{k'}$ .

Soient  $(X, \phi)$  et  $(X', \phi')$  deux systèmes dynamiques, tous deux à temps discret (respectivement à temps continu). Une *semi-conjugaison ensembliste* de  $(X, \phi)$  dans  $(X', \phi')$  est une application  $h : X \rightarrow X'$  surjective telle que

$$h \circ \phi = \phi' \circ h \quad (\text{respectivement } h \circ \phi^t = \phi'^t \circ h \text{ pour tout } t \in \mathbb{R}).$$

On dit que  $h$  est une *conjugaison ensembliste* si  $h$  est une bijection. Un système dynamique est *ensemblément semi-conjugué* (respectivement *ensemblément conjugué*) à un autre système dynamique s'il existe une semi-conjugaison (respectivement conjugaison) ensembliste du premier dans le second.

Dans le cas de deux systèmes dynamiques  $(X, \phi)$  et  $(X', \phi')$  clairement identifiés comme mesurables, mesurés, topologiques ou différentiables de classe  $C^k$ , une *semi-conjugaison* de  $(X, \phi)$  dans  $(X', \phi')$  est une semi-conjugaison ensembliste qui est de plus une application mesurable, préservant la mesure<sup>6</sup>, continue ou différentiable de classe  $C^k$  respectivement.

5. c'est-à-dire que  $\phi$  est continue et que pour toutes les cartes  $(U, \varphi)$  et  $(V, \psi)$  de l'atlas  $\mathcal{A}$  telles que  $\phi(\varphi(U)) \subset \psi(V)$ , l'application  $\psi^{-1} \circ \phi \circ \varphi : U \rightarrow V$  est de classe  $C^k$ .

6. Dans le cas de deux systèmes dynamiques mesurés, on demande que la semi-conjugaison soit seulement définie sur un ensemble invariant (voir la partie mesurable 0.2) de mesure totale, et on ne demande pas la condition de surjectivité de la semi-conjugaison, car la condition de préserver la mesure implique que l'image est de mesure totale.

Le système dynamique  $(X', \phi')$  à l'arrivée est alors appelé un *facteur* du premier. De même une *conjugaison* est une conjugaison ensembliste qui est de plus bimesurable, préservant la mesure ainsi que son inverse<sup>7</sup>, un homéomorphisme ou un  $C^k$ -difféomorphisme, respectivement. Deux systèmes dynamiques mesurables, mesurés, topologiques ou différentiables de classe  $C^k$  sont *conjugués* s'il existe une conjugaison entre eux. La relation « être conjugué à » est une relation d'équivalence sur tout ensemble de systèmes dynamiques de même type. En cas de doute, ou par exemple lorsque le degré de régularité de la (semi-)conjugaison est inférieur à celui des systèmes dynamiques, il est **fortement conseillé de préciser** quelle est la nature de la (semi-)conjugaison : mesurable, mesurée, topologique, de classe  $C^k$ , respectivement.

## 0.1 Orbites et orbites périodiques

Soient  $(X, \phi)$  un système dynamique et  $x$  un point de l'espace des phases  $X$ .

Si  $(X, \phi)$  est à temps discret, l'*orbite positive* de  $x$  (parfois juste appelée orbite), notée  $\mathcal{O}^+(x)$  (ou  $\mathcal{O}_\phi^+(x)$  voire  $\mathcal{O}_{(X,\phi)}^+(x)$  quand nous voulons préciser), est

$$\mathcal{O}^+(x) = \{\phi^n(x) : n \in \mathbb{N}\}$$

(où  $\phi^0 = \text{id}$  et par récurrence<sup>8</sup>  $\phi^{n+1} = \phi \circ \phi^n = \phi^n \circ \phi$  pour tout  $n \in \mathbb{N}$ ). Si  $(X, \phi)$  est à temps continu, l'*orbite positive* de  $x$ , notée  $\mathcal{O}^+(x)$  (ou  $\mathcal{O}_\phi^+(x)$  voire  $\mathcal{O}_{(X,\phi)}^+(x)$  quand nous voulons préciser), est

$$\mathcal{O}^+(x) = \{\phi^t(x) : t \in [0, +\infty[ \}.$$

Si  $(X, \phi)$  est à temps discret et inversible, l'*orbite* (parfois appelée *orbite complète*) de  $x$ , notée  $\mathcal{O}(x)$  (ou  $\mathcal{O}_\phi(x)$  voire  $\mathcal{O}_{(X,\phi)}(x)$  quand nous voulons préciser), est, en notant  $\phi^{-n} = (\phi^{-1})^n$  pour tout  $n \in \mathbb{N}$ ,

$$\mathcal{O}(x) = \{\phi^n(x) : n \in \mathbb{Z}\}.$$

Si  $(X, \phi)$  est à temps continu, l'*orbite* (parfois appelée *orbite complète*) de  $x$ , notée  $\mathcal{O}(x)$  (ou  $\mathcal{O}_\phi(x)$  voire  $\mathcal{O}_{(X,\phi)}(x)$  quand nous voulons préciser), est

$$\mathcal{O}(x) = \{\phi^t(x) : t \in \mathbb{R}\}.$$

Lorsque  $(X, \mathcal{B}, \mu, \phi)$  est un système dynamique mesuré, la théorie ergodique s'intéresse principalement au comportement de l'orbite positive (ou de l'orbite complète dans le cas inversible ou à temps continu) par  $\phi$  de  $\mu$ -presque tout point de  $X$ .

Si  $(X, \phi)$  est à temps discret, le point  $x$  est dit *périodique* s'il existe  $n \in \mathbb{N} - \{0\}$  tel que  $\phi^n(x) = x$ . Si  $(X, \phi)$  est à temps continu, le point  $x$  est dit *périodique* s'il existe  $t > 0$  tel que  $\phi^t(x) = x$ . Une *orbite périodique* est l'orbite positive d'un point périodique. Nous noterons  $\text{Per}(\phi)$  l'ensemble des points périodiques pour le système dynamique  $(X, \phi)$ .

La *période* d'un point périodique  $x$  d'un système dynamique à temps discret  $(X, \phi)$  est le minimum des  $n \in \mathbb{N} - \{0\}$  tels que  $\phi^n(x) = x$ , ou de manière équivalente, le cardinal de

7. On demande seulement que la conjugaison et son inverse soient définis sur un ensemble invariant (voir la partie mesurable 0.2) de mesure totale.

8. Lorsque  $X$  est un groupe multiplicatif et  $x \in X$ , le lecteur prendra bien garde à ne pas confondre  $\phi^n(x) = (\phi \circ \dots \circ \phi)(x)$  et  $\phi(x)^n = \phi(x) \times \dots \times \phi(x)$ .



son orbite positive. La *période* d'un point périodique  $x$  d'un système dynamique à temps continu  $(X, (\phi^t)_{t \in \mathbb{R}})$  est la borne inférieure<sup>9</sup> des  $t > 0$  tels que  $\phi^t(x) = x$ .

**Remarque.** Si  $h : X \rightarrow X'$  est une semi-conjugaison (ensembliste) entre deux systèmes dynamiques  $(X, \phi)$  et  $(X', \phi')$  de même type, alors pour tout temps  $t$  (dans  $\mathbb{N}$  s'ils sont à temps discret, dans  $\mathbb{Z}$  s'ils sont à temps discret et inversibles, et dans  $\mathbb{R}$  s'ils sont à temps continu), l'application  $h$  est aussi une semi-conjugaison (ensembliste) entre les systèmes dynamiques à temps discret  $(X, \phi^t)$  et  $(X', \phi'^t)$ . La semi-conjugaison  $h$  envoie point périodique sur point périodique :

$$h(\text{Per}(\phi)) \subset \text{Per}(\phi') , \quad (1)$$

envoie toute orbite positive de  $(X, \phi)$  dans une orbite positive de  $(X', \phi')$  :

$$\forall x \in X, \quad h(\mathcal{O}^+(x)) \subset \mathcal{O}^+(h(x)) , \quad (2)$$

et envoie toute orbite dans une orbite dans le cas à temps discret inversible, ou à temps continu :

$$\forall x \in X, \quad h(\mathcal{O}(x)) \subset \mathcal{O}(h(x)) . \quad (3)$$

De plus, si  $(X', \phi')$  est ensemblistement conjugué à  $(X, \phi)$  par une conjugaison ensembliste  $h : X \rightarrow X'$ , alors  $x$  est un point périodique de  $(X, \phi)$  si et seulement si  $h(x)$  est un point périodique de  $(X', \phi')$ , et les périodes de  $x$  et  $h(x)$  sont alors égales.

**Exemples.** (1) Dans ce texte, nous noterons  $\mathbb{S}_1 = \{z \in \mathbb{C} : |z| = 1\}$  le cercle unité du plan réel euclidien usuel. La *rotation* d'angle  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$  est l'homéomorphisme de  $\mathbb{S}_1$  dans  $\mathbb{S}_1$  défini par

$$z \mapsto e^{i\theta} z .$$

Elle est dite *rationnelle* si  $\theta \in 2\pi\mathbb{Q}$ , et *irrationnelle* sinon. Le système dynamique à temps discret  $(\mathbb{S}_1, z \mapsto e^{i\theta} z)$  est inversible. Si  $\theta = 2\pi\frac{p}{q}$  avec  $p \in \mathbb{Z}$ ,  $q \in \mathbb{N} - \{0\}$ , et  $p$  et  $q$  premiers entre eux, alors tout point de  $\mathbb{S}_1$  est périodique de période  $q$  pour la rotation d'angle  $\theta$ . Une rotation irrationnelle n'a pas d'orbite périodique, et toute orbite est dense.<sup>10</sup>

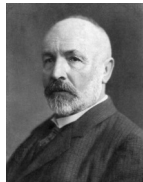
(2) Soit  $\mathcal{A}$  un ensemble non vide, muni de la topologie discrète, appelé un *alphabet*. Nous généraliserons cet exemple dans la partie 0.2. Le *système de Bernoulli*<sup>11</sup> *unilatère* (respectivement *bilatère*) d'alphabet  $\mathcal{A}$  est le système dynamique topologique à temps discret  $(\Sigma_+, \sigma_+)$  (respectivement  $(\Sigma, \sigma)$ ) où

9. Si  $(X, (\phi^t)_{t \in \mathbb{R}})$  est un système dynamique topologique et si  $X$  est séparé, alors cette borne inférieure est un minimum.

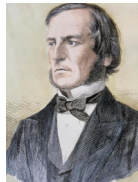
10. En effet, si  $\alpha = \frac{\theta}{2\pi} \notin \mathbb{Q}$ , alors pour tout  $t_0 \in \mathbb{R}$ , la partie  $t_0 + \mathbb{Z}\alpha + \mathbb{Z}$  est dense dans  $\mathbb{R}$  et s'envoie, par le revêtement (universel) de  $\mathbb{R}$  dans  $\mathbb{S}_1$  défini par  $t \mapsto e^{2i\pi t}$ , sur l'orbite de  $z_0 = e^{2i\pi t_0}$  par la rotation d'angle  $\theta$ .



Jacques Bernoulli  
1654-1705



Georg Cantor  
1845-1918



George Boole  
1815-1864



Constantin  
Carathéodory  
1873-1950



Eugene Dynkin  
1924-2014

11.

- l'espace des phases est l'espace topologique produit  $\Sigma_+ = \mathcal{A}^{\mathbb{N}}$  (respectivement  $\Sigma = \mathcal{A}^{\mathbb{Z}}$ ), dont les éléments  $(x_n)_{n \in \mathbb{N}}$  (respectivement  $(x_n)_{n \in \mathbb{Z}}$ ) sont appelés des *mots* infinis unilatères (respectivement bilatères) sur  $\mathcal{A}$ ,

- la transformation  $\sigma_+$  (respectivement  $\sigma$ ), appelée *décalage*<sup>12</sup> *de Bernoulli à gauche* est l'application qui à une suite  $(x_n)_{n \in \mathbb{N}}$  associe la suite  $(y_n)_{n \in \mathbb{N}}$  où  $y_n = x_{n+1}$  pour tout  $n$  dans  $\mathbb{N}$  (respectivement  $\mathbb{Z}$ ).

Le système de Bernoulli unilatère  $(\Sigma_+, \sigma_+)$  est non inversible si le cardinal de  $\mathcal{A}$  est au moins 2. Le système de Bernoulli bilatère  $(\Sigma, \sigma)$  est inversible. Le système de Bernoulli unilatère est un facteur du système de Bernoulli bilatère : l'application  $\pi_+ : \Sigma \rightarrow \Sigma_+$ , appelée *l'extension naturelle* (ou, de préférence, l'application d'oubli du passé), définie par  $(x_n)_{n \in \mathbb{Z}} \mapsto (x_n)_{n \in \mathbb{N}}$ , est une semi-conjugaison :

$$\pi_+ \circ \sigma = \sigma_+ \circ \pi_+ .$$

Puisque  $\mathcal{A}$  est discret, les espaces des phases  $\Sigma_+$  et  $\Sigma$  des systèmes de Bernoulli sont totalement discontinus<sup>13</sup> et métrisables respectivement pour les distances

$$\begin{aligned} \forall x = (x_n)_{n \in \mathbb{N}}, y = (y_n)_{n \in \mathbb{N}} \in \Sigma_+, \quad d(x, y) &= e^{-\sup \{ N \in \mathbb{N} : \forall n \in \llbracket 0, N \rrbracket, x_n = y_n \}} , \\ \forall x = (x_n)_{n \in \mathbb{Z}}, y = (y_n)_{n \in \mathbb{Z}} \in \Sigma, \quad d(x, y) &= e^{-\sup \{ N \in \mathbb{N} : \forall n \in \llbracket -N, N \rrbracket, x_n = y_n \}} , \end{aligned} \quad (4)$$

avec les conventions usuelles<sup>14</sup> que  $\sup \emptyset = 0$  et  $e^{-\infty} = 0$ . Lorsque l'alphabet  $\mathcal{A}$  est fini (ce qui sera en général le cas dans ces notes de cours), les espaces des phases  $\Sigma_+$  et  $\Sigma$  sont compacts, homéomorphes à un espace de Cantor<sup>11</sup>.<sup>15</sup>

Appelons *cylindre* (initial) du système de Bernoulli unilatère toute partie de  $\Sigma_+$  de la forme

$$[x_0, \dots, x_p] = \{ (y_n)_{n \in \mathbb{N}} \in \Sigma_+ : \forall n \in \llbracket 0, p \rrbracket, y_n = x_n \}$$

pour  $p$  dans  $\mathbb{N}$  (nous dirons que la *longueur* de ce cylindre est  $p + 1$ ) et  $x_0, \dots, x_p \in \mathcal{A}$ . L'ensemble des cylindres de  $(\Sigma_+, \sigma_+)$  est une base d'ouverts de  $\Sigma_+$ . Appelons *cylindre* (centré) du système de Bernoulli bilatère toute partie de  $\Sigma$  de la forme

$$[x_{-p}, \dots, x_p] = \{ (y_n)_{n \in \mathbb{Z}} \in \Sigma : \forall n \in \llbracket -p, p \rrbracket, y_n = x_n \}$$

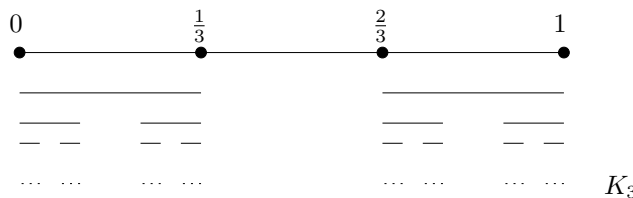
pour  $p \in \mathbb{N}$  (nous dirons que la *longueur* de ce cylindre est  $2p + 1$ ) et  $x_{-p}, \dots, x_p \in \mathcal{A}$ . L'ensemble des cylindres de  $(\Sigma, \sigma)$  est une base d'ouverts de  $\Sigma$ .

12. « shift » en anglais

13. Un espace topologique est *totalement discontinu* si ses composantes connexes sont des singletons. Tout produit d'espaces discrets (muni de la topologie produit) est totalement discontinu. Voir par exemple [Dug].

14. Ici,  $\emptyset$  est considérée comme une partie de  $\mathbb{N}$ , pas de  $\mathbb{R}$ , auquel cas la convention pourrait être différente.

15. Rappelons (voir par exemple [Dug]) qu'*un espace de Cantor* est un espace topologique compact métrisable totalement discontinu sans point isolé, et qu'ils sont tous homéomorphes à *l'espace de Cantor triadique*  $K_3 = \{ \sum_{i=1}^{\infty} \frac{a_i}{3^i} : a_i \in \{1, 2\} \}$ .



Les points périodiques de  $\sigma_+$  (respectivement  $\sigma$ ) sont les mots  $(x_n)_{n \in \mathbb{N}}$  (respectivement  $(x_n)_{n \in \mathbb{Z}}$ ) périodiques, c'est-à-dire tels qu'il existe  $k \in \mathbb{N} - \{0\}$  tel que  $x_{n+k} = x_n$  pour tout  $n$  dans  $\mathbb{N}$  (respectivement  $\mathbb{Z}$ ). Lorsque  $\mathcal{A}$  est fini, nous laissons au lecteur le soin de calculer le nombre de points périodiques de période (exactement)  $p \in \mathbb{N}$  donnée.

Les ensembles des points périodiques  $\text{Per}(\sigma_+)$  et  $\text{Per}(\sigma)$  sont denses dans  $\Sigma_+$  et  $\Sigma$  respectivement. En effet, dans le cas bilatère (l'argument est similaire dans le cas unilatère), pour tous les  $N \in \mathbb{N}$  et  $x = (x_n)_{n \in \mathbb{Z}} \in \Sigma$ , si  $y = (y_n)_{n \in \mathbb{Z}} \in \Sigma$  vérifie  $y_n = x_n$  pour  $-N \leq n \leq N$  et  $y_{n+2N+1} = y_n$  pour tout  $n \in \mathbb{Z}$ , alors  $y$  est périodique et à distance au plus  $e^{-N}$  de  $x$ .

## 0.2 Parties invariantes et mesures invariantes

La notion d'invariance pour les parties d'un espace des phases se décline en trois aspects.

a) Soit  $(X, \phi)$  un système dynamique. Une partie  $Y$  de l'espace des phases  $X$  est dite *positivement invariante* si  $\phi^t(Y) \subset Y$  pour tout temps  $t \geq 0$  (toujours avec la convention que  $t$  est entier si le temps est discret, et il suffit alors de le vérifier pour  $t = 1$ ).

Par exemple, une orbite positive est positivement invariante, et si  $(X, \phi)$  est un système dynamique topologique, alors l'adhérence  $\bar{Y}$  de toute partie positivement invariante  $Y$  l'est encore.<sup>16</sup>

b) Soient  $(X, \phi)$  un système dynamique topologique ou différentiable, et  $Y$  un sous-espace topologique ou une sous-variété de  $X$  respectivement, dont l'ensemble sous-jacent est une partie positivement invariante de  $X$ .

Si  $(X, \phi)$  est à temps discret, alors  $(Y, \phi|_Y)$  est un système dynamique topologique ou différentiable, à temps discret, appelé la *restriction* de  $(X, \phi)$  à  $Y$ .

Si  $(X, \phi)$  est à temps discret et inversible (respectivement s'il est à temps continu), alors  $Y$  est dite *invariante*, ou globalement invariante lorsqu'il convient d'insister, si  $\phi^t(Y) \subset Y$  pour tout temps  $t$ , ou, de manière équivalente, si  $\phi^t(Y) = Y$  pour tout temps  $t$ . Ceci implique que  $(Y, \phi|_Y)$  (respectivement  $(Y, (\phi^t|_Y)_{t \in \mathbb{R}})$ ) est un système dynamique topologique ou différentiable, à temps discret inversible (respectivement à temps continu), appelé la *restriction* de  $(X, \phi)$  à  $Y$ . Par exemple, si  $A$  est une partie de  $X$ , alors l'intersection  $\bigcap_{t \in \mathbb{Z}} \phi^t(A)$  en temps discret et  $\bigcap_{t \in \mathbb{R}} \phi^t(A)$  en temps continu est un sous-espace topologique<sup>17</sup> et globalement invariant.

c) Soit  $(X, \phi)$  un système dynamique mesurable (respectivement mesuré) à temps discret. Une partie mesurable  $Y$  de  $X$  est dite *invariante*<sup>18</sup> si  $\phi^{-1}(Y) = Y$ . Cette condition implique la condition d'être positivement invariante, mais elle est en général bien plus forte que la condition d'être positivement invariante. Par exemple, si  $A$  est une partie mesurable, alors les parties suivantes sont mesurables et invariantes

- $\limsup_{n \in \mathbb{N}} \phi^{-n}(A) = \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} \phi^{-k}(A)$  des éléments  $x \in X$  tels qu'il existe une infinité de  $n \in \mathbb{N}$  pour lesquels  $\phi^n(x) \in A$ ,

16. En effet, en temps discret (la démonstration est analogue en temps continu), soit  $x \in \bar{Y}$ . Pour tout voisinage ouvert  $U$  de  $\phi(x)$ , montrons que  $U$  rencontre  $Y$ , ce qui implique que  $\phi(x) \in \bar{Y}$ . L'ouvert  $\phi^{-1}(U)$ , qui contient  $x$ , rencontre  $Y$ . Si  $y \in \phi^{-1}(U) \cap Y$ , alors  $U \cap Y$  contient  $\phi(y)$ , qui appartient à  $Y$ .

17. Même lorsque  $(X, \phi)$  est un système dynamique différentiable et  $A$  une sous-variété, la partie  $\bigcap_{t \in \mathbb{Z}} \phi^t(A)$  ou  $\bigcap_{t \in \mathbb{R}} \phi^t(A)$  n'est pas forcément une sous-variété, voir par exemple la partie 10.5

18. Bien noter la différence sémantique entre le cas des systèmes dynamiques topologiques ou différentiables et celui des systèmes dynamiques mesurables ou mesurés.

•  $\liminf_{n \in \mathbb{N}} \phi^{-n}(A) = \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} \phi^{-k}(A)$  des éléments  $x \in X$ , tels que pour tout  $n \in \mathbb{N}$  assez grand, nous ayons  $\phi^n(x) \in A$ .

Notons  $\mathcal{B}$  la tribu (respectivement  $\mathcal{B}$  et  $\mu$  la tribu et la mesure) de  $(X, \phi)$ . Si une partie invariante  $Y$  est munie de la tribu induite<sup>19</sup>  $\mathcal{B}_Y$  (respectivement de la tribu induite  $\mathcal{B}_Y$  et de la mesure restreinte<sup>20</sup>  $\mu|_{\mathcal{B}_Y}$  de  $\mu$  à  $\mathcal{B}_Y$ , alors  $(Y, \phi|_Y)$  est un système dynamique mesurable (respectivement mesuré), appelé la *restriction* de  $(X, \phi)$  à  $Y$ .

Si  $(X, \phi)$  est un système dynamique mesuré à temps discret, et si  $\mu$  est sa mesure, une partie mesurable  $Y$  de  $X$  est dite *presque invariante*<sup>21</sup> si  $\mu(\phi^{-1}(Y)\Delta Y) = 0$ , où  $\Delta$  désigne la *différence symétrique* d'ensembles, définie par  $A\Delta B = (A \setminus B) \cup (B \setminus A)$ .<sup>22</sup>

Si  $(X, \phi)$  est un système dynamique et  $f : X \rightarrow \mathbb{C}$  une application, nous dirons que  $f$  est *invariante*<sup>23</sup> si  $f \circ \phi = f$  en temps discret et si  $f \circ \phi^t = f$  pour tout  $t \in \mathbb{R}$  en temps continu. Nous dirons dans les deux cas que  $f \circ \phi = f$ . Par exemple, si  $f = \mathbb{1}_A$  est la fonction caractéristique d'une partie  $A$  de  $X$ , et si  $(X, \phi)$  est à temps discret et inversible ou à temps continu, alors dire que  $f$  est invariante est équivalent à dire que  $A$  est globalement invariante. Si  $(X, \mu, \phi)$  est un système dynamique mesuré, nous dirons que  $f$  est *presque invariante*<sup>24</sup> si  $f \circ \phi = f$  presque partout en temps discret et si pour tout  $t \in \mathbb{R}$ , nous avons l'égalité  $f \circ \phi^t = f$  presque partout. Nous dirons dans les deux cas que  $f \circ \phi = f$  presque partout.

Développons maintenant la notion de mesure invariante (voir le début de la partie 0 pour des définitions). Si  $(X, \mathcal{B}, \phi)$  est un système dynamique mesurable, et si  $\mu$  est une mesure sur  $(X, \mathcal{B})$ , rappelons que  $\mu$  est invariante par ce système dynamique<sup>25</sup> si pour tout temps  $t$  (dans  $\mathbb{N}$  ou  $\mathbb{Z}$  en temps discret (il suffit alors de vérifier pour  $t = 1$ ) et dans  $\mathbb{R}$  en temps continu), nous avons

$$(\phi^t)_*\mu = \mu.$$

Par la définition de l'intégrale des fonctions positives mesurables et celle des fonctions intégrables (voir par exemple [Coh, §2.3]), la mesure  $\mu$  est invariante si et seulement si pour tout temps  $t$  (il suffit de vérifier pour  $t = 1$  en temps discret) et pour toute fonction mesurable  $f : X \rightarrow [0, +\infty[$ , nous avons

$$\int_X f \circ \phi^t d\mu = \int_X f d\mu.$$

et si et seulement si pour tout temps  $t$ , nous avons

$$\forall f \in \mathbb{L}^1(X, \mu), \quad \int_X f \circ \phi^t d\mu = \int_X f d\mu.$$

Par la linéarité et la continuité en  $f \in \mathbb{L}^1(X, \mu)$  de l'égalité ci-dessus, la mesure  $\mu$  est invariante si et seulement s'il existe une partie  $\mathcal{E}$  de  $\mathbb{L}^1(X, \mu)$  engendrant un sous-espace

19. l'ensemble  $\mathcal{B}_Y = \{Y \cap B : B \in \mathcal{B}\}$  des intersections avec  $Y$  des parties mesurables de  $X$

20. Il convient de ne pas confondre la mesure restreinte  $\mu|_{\mathcal{B}_Y}$  avec la mesure induite (voir l'exercice E.13)  $\mu_Y$  de  $\mu$  sur  $Y$ , celle-ci n'étant définie que lorsque la mesure de  $Y$  est finie non nulle par  $\mu_Y = \frac{1}{\mu(Y)}\mu|_{\mathcal{B}_Y}$ .

21. ou  *$\mu$ -presque invariante* lorsqu'il convient de préciser

22. Remarquons que  $\mu(A\Delta B) = \|\mathbb{1}_A - \mathbb{1}_B\|_{\mathbb{L}^1}$ .

23. ou  *$\phi$ -invariante* lorsqu'il convient de préciser

24. ou  *$\mu$ -presque partout  $\phi$ -invariante* lorsqu'il convient de préciser

25. ou  *$\phi$ -invariante* lorsqu'il convient de préciser

vecotirel dense de  $\mathbb{L}^1(X, \mu)$  tel que pour tout temps  $t$  (il suffit de vérifier pour  $t = 1$  en temps discret) nous avons

$$\forall f \in \mathcal{E}, \quad \int_X f \circ \phi^t d\mu = \int_X f d\mu. \quad (5)$$

Par exemple, la mesure de comptage sur une orbite périodique d'un système dynamique à temps discret est une mesure invariante. Nous construirons dans la partie 6 de nombreuses mesures invariantes sur les systèmes de Bernoulli, autres que ces mesures de comptages sur les orbites périodiques.

Donnons des critères pratiques, qui essaient de minimiser le nombre de vérifications, pour montrer qu'une application  $\phi$  préserve la mesure  $\mu$ . Ceci signifie a priori de vérifier que  $(\phi_*\mu)(B) = \mu(B)$  pour tout  $B$  dans la tribu, en général non dénombrable. Restreindre ce nombre de vérifications est fondamental.<sup>26</sup>

Soit  $X$  un ensemble. Une *algèbre de Boole*<sup>11</sup> de parties de  $X$  est une partie  $\mathcal{E}$  de l'ensemble  $\mathcal{P}(X)$  des parties de  $E$ , contenant  $\emptyset$ , telle que l'intersection de deux éléments quelconques de  $\mathcal{E}$ , et le complémentaire de tout élément de  $\mathcal{E}$ , soient des éléments de  $\mathcal{E}$ . En particulier, nous avons  $X \in \mathcal{E}$ .

Un *semi-anneau de Boole* sur  $X$  est une partie  $\mathcal{E}$  de l'ensemble  $\mathcal{P}(X)$  de toutes les parties de  $X$ , contenant  $\emptyset$ , telle que l'intersection de deux éléments de  $\mathcal{E}$  soit un élément de  $\mathcal{E}$ , et la différence  $A \setminus B = \{x \in A : x \notin B\}$  de deux éléments  $A$  et  $B$  de  $\mathcal{E}$  soit une union finie d'éléments de  $\mathcal{E}$ . Par exemple, une algèbre de Boole est un semi-anneau de Boole, et l'ensemble des intervalles semi-ouverts de  $\mathbb{R}$  est un semi-anneau de Boole, mais pas une algèbre de Boole.

Un  $\pi$ -système sur  $X$  est une partie  $\mathcal{E}$  de l'ensemble  $\mathcal{P}(X)$  de toutes les parties de  $X$ , qui contient  $X$  et telle que l'intersection de deux éléments de  $\mathcal{E}$  appartienne à  $\mathcal{E}$ . Par exemple, un semi-anneau de Boole est un  $\pi$ -système. Nous verrons un exemple important à la fin de cette partie 0.2.

Le résultat suivant, qui utilise le procédé de Carathéodory<sup>11</sup> de construction de mesures par les mesures extérieures, est contenu par exemple dans [AIB, Theo. 9.22] pour l'assertion (1) et dans [Coh], [Neve, page 23] pour l'assertion (2). Nous donnons deux versions, car en fonction des exemples étudiés, il n'est pas toujours facile de vérifier la propriété de  $\sigma$ -additivité sur le semi-anneau de Boole qu'il est pourtant naturel de considérer.

### **Théorème 0.1. (Théorème d'extension de Carathéodory)**

- (1) Soient  $\mathcal{E}$  un semi-anneau de Boole sur un ensemble  $X$ ,  $\mathcal{B}$  la  $\sigma$ -algèbre engendrée par  $\mathcal{E}$  et  $\mu : \mathcal{E} \rightarrow [0, +\infty]$  une application vérifiant :

**( $\sigma$ -additivité sur  $\mathcal{E}$ )** pour toute suite  $(C_n)_{n \in \mathbb{N}}$  d'éléments de  $\mathcal{E}$  deux à deux disjoints et de réunion appartenant à  $\mathcal{E}$ , nous avons  $\mu(\bigcup_{n \in \mathbb{N}} C_n) = \sum_{n \in \mathbb{N}} \mu(C_n)$ .

Alors  $\mu$  se prolonge en une mesure (encore notée  $\mu$ ) sur  $\mathcal{B}$ . De plus, si  $\mu$  est  $\sigma$ -finie sur  $\mathcal{E}$  (c'est-à-dire si  $X$  est une union dénombrable d'éléments de  $\mathcal{E}$  de mesure finie pour  $\mu$ ), alors cette extension est  $\sigma$ -finie et unique.

26. Une autre approche (voir la partie 0.3), dite fonctionnelle, dans le cas où l'espace des phases est un espace topologique localement compact à base dénombrable d'ouvert, est de minimiser le nombre de vérifications que pour toute fonction réelle  $f \in C_0^0(X; \mathbb{R})$  sur  $X$ , continue et s'annulant à l'infini, nous avons  $\int_X f \circ \phi d\mu = \int_X f d\mu$ , en ne vérifiant cette égalité que pour une famille aussi petite que possible de telles fonctions  $f$ .

- (2) Soient  $\mathcal{E}$  une algèbre de Boole de parties d'un ensemble  $X$  et  $\mathcal{B}$  la  $\sigma$ -algèbre de parties de  $X$  engendrée par  $\mathcal{E}$ . Considérons  $\mu : \mathcal{E} \rightarrow [0, +\infty]$  une application vérifiant
- i) l'application  $\mu$  est  $\sigma$ -finie sur  $\mathcal{E}$  : l'ensemble  $X$  est une union dénombrable d'éléments de  $\mathcal{E}$  dont l'image par  $\mu$  est finie,
  - ii) l'application  $\mu$  est additive sur  $\mathcal{E}$  : pour tous les éléments  $C_1, C_2$  de  $\mathcal{E}$  d'intersection  $C_1 \cap C_2$  vide, nous avons l'égalité  $\mu(C_1 \cup C_2) = \mu(C_1) + \mu(C_2)$ ,
  - iii) l'application  $\mu$  est continue sur  $\mathcal{E}$  : pour toute suite décroissante (pour l'inclusion)  $(C_n)_{n \in \mathbb{N}}$  d'éléments de  $\mathcal{E}$  telle que  $\mu(C_0) < \infty$  et  $\bigcap_{n \in \mathbb{N}} C_n = \emptyset$ , nous avons  $\lim_{n \rightarrow \infty} \mu(C_n) = 0$ .

Alors  $\mu$  se prolonge de manière unique en une mesure  $\sigma$ -finie sur  $\mathcal{B}$ .  $\square$

En particulier, si  $X \in \mathcal{E}$  et si l'application  $\mu : \mathcal{E} \rightarrow [0, +\infty]$  comme dans l'énoncé (1) ou (2) ci-dessus vérifie de plus  $\mu(X) = 1$ , alors son extension  $\mu$  est unique et c'est une mesure de probabilité. Le premier critère ci-dessous est une conséquence de l'unicité dans l'assertion (1) du théorème d'extension de Carathéodory ci-dessus. Le second critère d'égalité ci-dessous utilise les arguments de classe monotone de Dynkin,<sup>11</sup> voir [Coh, Coro. 1.6.3].

**Proposition 0.2.** Soit  $(X, \mathcal{B})$  un espace mesurable.

- (1) Deux mesures  $\sigma$ -finies sur  $(X, \mathcal{B})$ , qui coïncident sur un sous-anneau de Boole (et en particulier sur une sous-algèbre de Boole) de  $\mathcal{B}$  qui engendrent la  $\sigma$ -algèbre  $\mathcal{B}$ , sont égales.
- (2) Deux mesures finies sur  $(X, \mathcal{B})$ , qui coïncident sur  $\{X\}$  et sur un  $\pi$ -système constitué d'éléments de  $\mathcal{B}$  engendrant la  $\sigma$ -algèbre  $\mathcal{B}$ , sont égales.
- (3) Soient  $(X, \mathcal{A}, \mu)$  et  $(Y, \mathcal{B}, \nu)$  deux espaces mesurés et  $\phi : X \rightarrow Y$  une application mesurable. Soit  $\mathcal{C}$  une semi-algèbre de Boole dans  $\mathcal{B}$  (i.e. une partie de  $\mathcal{B}$ , contenant  $\emptyset$ , telle que l'intersection de deux éléments de  $\mathcal{C}$ , et le complémentaire d'un élément de  $\mathcal{C}$ , soient une union finie disjointe d'éléments de  $\mathcal{C}$ ), qui engendrent la  $\sigma$ -algèbre  $\mathcal{B}$ . Si  $\nu$  est  $\sigma$ -finie sur  $\mathcal{C}$  et si  $\phi$  préserve la mesure sur  $\mathcal{C}$ , i.e. si

$$\forall B \in \mathcal{C}, \quad \phi^{-1}(B) \in \mathcal{A} \quad \text{et} \quad \mu(\phi^{-1}(B)) = \nu(B),$$

alors  $\phi$  préserve la mesure.

**Démonstration.** (3) Soit  $\mathcal{E}$  l'ensemble des unions finies disjointes d'éléments de  $\mathcal{C}$ . Alors  $\mathcal{E}$  est une algèbre de Boole, et les mesures  $\phi_*\mu$  et  $\nu$  coïncident sur  $\mathcal{E}$ . Par l'unicité dans le théorème 0.1 (2), ces mesures sont égales, donc  $\phi$  préserve la mesure.  $\square$

**Exemples.** Nous généralisons maintenant les exemples (1) et (2) de la partie 0.1, en oubliant la topologie, mais en rajoutant des mesures invariantes. Nous donnons aussi des exemples de mesures invariantes venant de la physique, en particulier de la théorie des gaz parfaits.

(1) (Systèmes de Kronecker<sup>27</sup>)



27. Leopold Kronecker  
1823-1891



Alfred Haar  
1885-1933



Henri Lebesgue  
1875-1941



Joseph Liouville  
1809-1882



Augustin Cauchy  
1789-1857

Soit  $G$  un groupe topologique<sup>28</sup> localement compact. Il existe une mesure  $\nu_G$  sur  $G$  invariante par toutes les translations à gauche  $\lambda_g : x \mapsto gx$  où  $g \in G$ , unique modulo multiplication par un réel strictement positif, appelée la *mesure de Haar*<sup>27</sup>, voir par exemple [Wei]. Par exemple, la mesure de Lebesgue<sup>27</sup> sur  $\mathbb{R}$  est l'unique mesure de Haar  $\nu_G$  de  $G = (\mathbb{R}, +)$  telle que  $\nu_G([0, 1]) = 1$ .

Lorsque  $G$  est compact, la *mesure de Haar normalisée*  $\nu_G$  de  $G$  est l'unique mesure de Haar de  $G$  qui est une mesure de probabilité. Par exemple, si  $G = (\mathbb{S}_1, \times)$ , alors l'unicité et l'invariance de la mesure de Lebesgue de  $\mathbb{R}$  par translations montre que la mesure de Haar normalisée  $\nu_G$  de  $G$  est déterminée par

$$\forall f \in C^0(\mathbb{S}_1; \mathbb{R}), \quad \nu_G(f) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta}) d\theta.$$

Nous appellerons *système de Kronecker* tout système dynamique topologique  $(G, \lambda_g)$  ou tout système dynamique mesuré  $(G, \nu_G, \lambda_g)$  où  $g \in G$ . Nous étudierons les propriétés ergodiques de tels systèmes lorsque  $G = \mathrm{SL}_N(\mathbb{R})$  dans la partie 9.

(2) **(Variables aléatoires indépendantes identiquement distribuées.)**

Soit  $(\Omega, \mathcal{B}, \mu)$  un espace de probabilité. L'ensemble produit  $\Omega^{\mathbb{N}} = \{(\omega_n)_{n \in \mathbb{N}} : \omega_n \in \Omega\}$  est muni d'une transformation  $\sigma_+ : \Omega^{\mathbb{N}} \rightarrow \Omega^{\mathbb{N}}$ , appelée le *décalage* (à gauche), qui à une suite  $(\omega_n)_{n \in \mathbb{N}}$  associe la suite  $(\omega'_n)_{n \in \mathbb{N}}$  où  $\omega'_n = \omega_{n+1}$  pour tout  $n$  dans  $\mathbb{N}$ . Un *cylindre* (initial) de l'ensemble produit est une partie de  $\Omega^{\mathbb{N}}$  de la forme

$$[\omega_0 \in E_0, \dots, \omega_k \in E_k] = \{(\omega_n)_{n \in \mathbb{N}} \in \Omega^{\mathbb{N}} : \forall i \in \llbracket 0, k \rrbracket, \omega_i \in E_i\},$$

où  $k \in \mathbb{N}$ , et  $E_0, \dots, E_k \in \mathcal{B}$ . La *tribu produit*  $\mathcal{B}^{\mathbb{N}}$  sur l'ensemble produit  $\Omega^{\mathbb{N}}$  est la tribu engendrée par les cylindres. La *mesure de Bernoulli*  $\mu^{\mathbb{N}}$  est l'unique mesure de probabilité sur l'espace mesurable  $(\Omega^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}})$  telle que pour tous les  $k \in \mathbb{N}$ , et  $E_0, \dots, E_k \in \mathcal{B}$ , nous ayons

$$\mu^{\mathbb{N}}[\omega_0 \in E_0, \dots, \omega_k \in E_k] = \prod_{i=0}^k \mu(E_i); \quad (6)$$

elle est invariante par le décalage  $\sigma_+$ . Son existence, son unicité et son invariance découlent du théorème de Carathéodory 0.1 (2) et de la proposition 0.2 (3). En effet, l'ensemble  $\mathcal{E}$  des unions finies disjointes de cylindres initiaux de l'ensemble produit  $\Omega^{\mathbb{N}}$  est une algèbre de Boole<sup>29</sup> engendrant par définition la tribu produit  $\mathcal{B}^{\mathbb{N}}$ . Cette algèbre de Boole est

---

28. Un *groupe topologique* est un groupe  $G$  muni d'une topologie qui rend continues les applications de multiplication et de passage à l'inverse, ou, de manière équivalente, qui rend continue l'application  $(x, y) \mapsto xy^{-1}$ . Par exemple, si  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{K} = \mathbb{C}$ , en munissant tout espace vectoriel sur  $\mathbb{K}$  de dimension finie de l'unique topologie définie par n'importe quelle norme (elles sont toutes équivalentes), tout sous-groupe de  $\mathrm{GL}_N(\mathbb{K})$  muni de la topologie induite par celle de l'espace vectoriel  $\mathcal{M}_N(\mathbb{K})$ , est un groupe topologique, car les opérations de multiplication des matrices et de passage à l'inverse sont respectivement polynomiales et rationnelles (de dénominateur ne s'annulant pas) en les coefficients. Un *morphisme de groupes topologiques* entre deux groupes topologiques est un morphisme de groupes continu.

Pour tout  $g$  dans  $G$ , nous noterons  $\lambda_g$  et  $\rho_g$  les homéomorphismes de  $G$  définis par  $\lambda_g : x \mapsto gx$  (*translation à gauche*) et  $\rho_g : x \mapsto xg^{-1}$  (*translation à droite*, certains ouvrages utilisent  $x \mapsto xg$ ).

29. En effet, pour tous les  $k, k' \in \mathbb{N}$  tels que  $k \leq k'$  (par symétrie) et  $E_0, \dots, E_k, E'_0, \dots, E'_{k'} \in \mathcal{B}$ , nous avons

$$\begin{aligned} & [\omega_0 \in E_0, \dots, \omega_k \in E_k] \cap [\omega_0 \in E'_0, \dots, \omega_{k'} \in E'_{k'}] \\ &= [\omega_0 \in E_0 \cap E'_0, \dots, \omega_k \in E_k \cap E'_k, \omega_{k+1} \in E'_{k+1}, \dots, \omega_{k'} \in E'_{k'}] \end{aligned}$$

invariante par le décalage, car  $\sigma_+([\omega_0 \in E_0, \dots, \omega_k \in E_k]) = [\omega_0 \in E_1, \dots, \omega_{k-1} \in E_k]$ . Considérons l'application  $\mu^{\mathbb{N}} : \mathcal{E} \rightarrow [0, +\infty[$  définie par les formules (6) sur l'ensemble des cylindres et étendue par additivité sur les unions finies disjointes de cylindres.

**Lemme 0.3.** *Cette application est bien définie : elle ne dépend pas de la manière d'écrire les éléments de  $\mathcal{E}$  comme union disjointe de cylindres.*

**Démonstration.** Montrons par récurrence sur  $n \in \mathbb{N}$  et par subdivision commune que si  $(C_i)_{i \in I}$  et  $(C'_j)_{j \in J}$  sont deux familles finies de cylindres deux à deux disjoints de longueur au plus  $n$  ayant la même réunion  $C$ , alors il existe une famille finie  $(C''_k)_{k \in K}$  de cylindres deux à deux disjoints de réunion  $C$  telle que  $\sum_{i \in I} \mu(C_i) = \sum_{k \in K} \mu(C''_k) = \sum_{j \in J} \mu(C'_j)$ , ce qui montre le résultat. Pour simplifier, pour tous les  $k \in \mathbb{N}$  et  $B \in \mathcal{B}$ , nous noterons  $[E]_k = [\omega_0 \in X, \dots, \omega_{k-1} \in X, \omega_k \in E]$ .

Le résultat est vrai pour  $n = 0$ , car alors en écrivant  $C_i = [E_i]_0$  et  $C'_j = [E'_j]_0$ , nous avons  $\bigcup_{i \in I} E_i = \bigcup_{j \in J} E'_j$  et ces réunions sont disjointes dans  $X$ , donc le résultat en découle par l'additivité de  $\mu$  en considérant  $(C''_{ij} = [E_i \cap E'_j]_0)_{(i,j) \in I \times J}$ .

Supposons le résultat vrai pour  $n \in \mathbb{N}$ . Écrivons  $C_i = D_i \cap [E_i]_{n+1}$  et  $C'_j = D'_j \cap [E'_j]_{n+1}$  avec  $D_i$  et  $D'_j$  des cylindres de longueur au plus  $n$ . Par subdivision commune, il existe une famille finie  $(E''_k)_{k \in K}$  de boréliens deux à deux disjoints, que nous pouvons supposer non vides, et pour tout  $i \in I$  une partie  $K_i$  de  $K$  et pour tout  $j \in J$  une partie  $K'_j$  de  $K$  telles que  $E_i = \bigcup_{k \in K_i} E''_k$  et  $E'_j = \bigcup_{k \in K'_j} E''_k$ . Par l'additivité en la partie  $E_k$  de la formule (6), nous avons

$$\sum_{i \in I} \mu^{\mathbb{N}}(C_i) = \sum_{i \in I} \sum_{k \in K_i} \mu^{\mathbb{N}}(D_i \cap [E''_k]_{n+1}) = \sum_{k \in K} \left( \sum_{i \in I: k \in K_i} \mu^{\mathbb{N}}(D_i) \right) \mu(E''_k)$$

et

$$\sum_{j \in J} \mu^{\mathbb{N}}(C'_j) = \sum_{j \in J} \sum_{k \in K'_j} \mu^{\mathbb{N}}(D'_j \cap [E''_k]_{n+1}) = \sum_{k \in K} \left( \sum_{j \in J: k \in K'_j} \mu^{\mathbb{N}}(D'_j) \right) \mu(E''_k).$$

Pour tout  $k \in K$ , puisque les  $E''_k$  sont non vides et deux à deux disjoints, les familles  $(D_i)_{i \in I: k \in K_i}$  et  $(D'_j)_{j \in J: k \in K'_j}$  sont deux familles finies de cylindres deux à deux disjoints de longueur au plus  $n$  ayant la même réunion. Le résultat en découle donc par l'hypothèse de récurrence.  $\square$

L'application  $\mu^{\mathbb{N}} : \mathcal{E} \rightarrow [0, +\infty[$  vérifie

- la propriété **i)** du théorème 0.1 (2) car  $X = [\omega_0 \in X] \in \mathcal{E}$  et  $\mu^{\mathbb{N}}(X) = 1$ ,
- la propriété **ii)** par construction et
- la propriété **iii)**. En effet, conservons pour tous les  $k \in \mathbb{N}$  et  $B \in \mathcal{B}$  la notation  $[E]_k = [\omega_0 \in X, \dots, \omega_{k-1} \in X, \omega_k \in E]$ . Rappelons qu'un produit infini (dénombrable pour n'utiliser que l'axiome du choix dénombrable) est non vide si et seulement si chaque facteur est non vide. Toute suite décroissante  $(C_j)_{j \in \mathbb{N}}$  de cylindres non vides s'écrit de la

et

$${}^c[\omega_0 \in E_0, \dots, \omega_k \in E_k] = [\omega_0 \in {}^c E_0] \sqcup \bigcap_{i=1}^k [\omega_0 \in E_0, \dots, \omega_{i-1} \in E_{i-1}, \omega_i \in {}^c E_i],$$

donc l'intersection de deux cylindres est un cylindre (et en particulier l'ensemble des cylindres est un  $\pi$ -système), et le complémentaire d'un cylindre est une union finie disjointe de cylindres. Le résultat découle alors des formules ensemblistes  ${}^c(\bigcup_{i=1}^k A_i) = \bigcap_{i=1}^k {}^c A_i$ ,  $(\bigcup_{i=1}^k A_i) \cap (\bigcup_{j=1}^{\ell} B_j) = \bigcup_{i=1}^k (A_i \cap B_j)$ . Notons que comme  $A \cup B = (A \cap {}^c B) \sqcup (B \cap {}^c A) \sqcup (A \cap B)$ , toute union finie de cylindres est une union finie disjointe de cylindres.



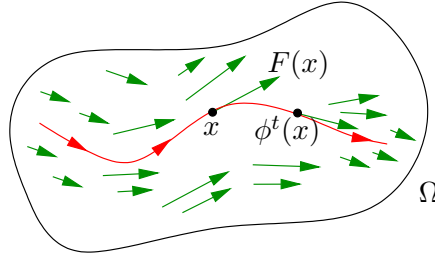
manière suivante. Pour tout  $j \in \mathbb{N}$ , il existe une suite  $(B_{i,j})_{i \in \mathbb{N}}$  dans  $\mathcal{B}$  telle que  $B_{i,j} = X$  pour tout  $i \in \mathbb{N}$  assez grand et  $C_j = \bigcap_{i \in \mathbb{N}} [B_{i,j}]_i$ , et telle que pour tout  $i \in \mathbb{N}$ , nous avons  $B_{i,j+1} \subset B_{i,j}$ . Si  $\bigcap_{j \in \mathbb{N}} C_j$  est vide, alors il existe  $i_0 \in \mathbb{N}$  tel que  $\bigcap_{j \in \mathbb{N}} B_{i_0,j}$  est vide. Par les propriétés des mesures finies, nous avons donc  $\lim_{j \rightarrow +\infty} \mu(B_{i_0,j}) = 0$ . Nous avons donc comme souhaité

$$0 \leq \limsup_{j \rightarrow +\infty} \mu^{\mathbb{N}}(C_j) = \limsup_{j \rightarrow +\infty} \prod_{i \in \mathbb{N}} \mu(B_{i,j}) \leq \limsup_{j \rightarrow +\infty} \mu(B_{i_0,j}) = 0.$$

Le quadruplet  $(\Omega^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}}, \mu^{\mathbb{N}}, \sigma_+)$  est donc un système dynamique mesuré.<sup>30</sup>

**Exercice E.1.** <sup>31</sup> Soient  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires indépendantes identiquement distribuées, définies sur un espace de probabilité  $(\tilde{\Omega}, \tilde{\mathcal{B}}, \mathbb{P})$  et à valeurs dans un espace mesurable  $(\Omega, \mathcal{B})$ , et  $\mu$  est la loi commune.<sup>32</sup> Montrer qu'il existe une application préservant la mesure  $\Theta$  de  $(\tilde{\Omega}, \tilde{\mathcal{B}}, \mathbb{P})$  dans  $(\Omega^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}}, \mu^{\mathbb{N}})$  telle que pour tout  $k \in \mathbb{N}$ , si  $\text{pr}_k : \Omega^{\mathbb{N}} \rightarrow \Omega$  est la  $k$ -ème projection  $(x_n)_{n \in \mathbb{N}} \mapsto x_k$ , alors  $X_k = \text{pr}_k \circ \Theta$ .

(3) (Mesures de Liouville<sup>27</sup> et systèmes hamiltoniens.)



Soient  $\Omega$  un ouvert non vide de  $\mathbb{R}^n$  et  $F : \Omega \rightarrow \mathbb{R}^n$  un champ de vecteurs de classe  $C^\infty$  sur  $\Omega$ . Considérons l'équation différentielle ordinaire du premier ordre

$$\frac{dy}{dt} = F(y)$$

sur  $\Omega$ . Notons  $\phi^t(x)$  la valeur en l'instant  $t$  de l'unique solution valant  $x$  à l'instant  $t = 0$  lorsqu'elle est définie. Supposons que le champ de vecteurs  $F$  soit *complet*, c'est-à-dire que

30. Le système de Bernoulli unilatère construit dans la partie 0.1 correspond au cas où  $\mathcal{B} = \mathcal{P}(\Omega)$  est la  $\sigma$ -algèbre de la topologie discrète sur l'alphabet  $\Omega$ , mais il est considéré comme un système dynamique topologique, et ne vient pas avec une mesure invariante particulière. Les cylindres au sens de la partie 0.1 sont les cylindres  $[\omega_0 \in E_0, \dots, \omega_{k-1} \in E_{k-1}, \omega_k \in E_k]$  au sens de cette partie où les  $E_0, \dots, E_k$  sont des singletons. Notons que l'ensemble des cylindres au sens de la partie 0.1, auquel nous rajoutons l'ensemble vide, est un sous-anneau de Boole si et seulement si l'alphabet  $\mathcal{A}$  est fini.

31. Voir la partie 0.5 pour des indications de solutions.

32. Une *variable aléatoire*  $v$  définie sur  $(\tilde{\Omega}, \tilde{\mathcal{B}}, \mathbb{P})$  à valeurs dans  $(\Omega, \mathcal{B})$  est une application mesurable de  $\tilde{\Omega}$  dans  $\Omega$ . Sa *loi* est la mesure image  $v_*\mathbb{P}$ . Son *espérance* est  $\mathbb{E}[v] = \int_{\tilde{\Omega}} v \, d\mathbb{P}$ . Une suite  $(v_i)_{i \in \mathbb{N}}$  de variables aléatoires est *uniformément distribuée* si les lois des  $v_i$  coïncident, et *indépendante* si pour toute partie finie  $F$  de  $\mathbb{N}$ , pour toute famille finie  $(A_i)_{i \in F}$  dans  $\mathcal{B}$ , nous avons

$$\mathbb{P}\left(\bigcap_{i \in F} v_i^{-1}(A_i)\right) = \prod_{i \in F} \mathbb{P}(v_i^{-1}(A_i)).$$

$\phi^t(x)$  est défini pour tous les  $t \in \mathbb{R}$  et  $x \in \Omega$ . Alors le théorème de Cauchy<sup>27</sup>-Lipshitz<sup>33</sup> dit que  $(\Omega, \phi = (\phi^t)_{t \in \mathbb{R}})$  est un système dynamique différentiable de classe  $C^\infty$  à temps continu, dont les orbites sont les courbes intégrales du champ de vecteurs.

Le résultat suivant donne un critère d'existence d'une mesure invariante pour le système dynamique  $(\Omega, \phi)$  qui est absolument continue par rapport à la mesure de Lebesgue sur  $\Omega$ . Une telle mesure est appelée une *mesure de Liouville*. Il reste valable en remplaçant le flot de  $F$  par le flot local (maximal) de  $F$  si le champ de vecteurs  $F$  n'est pas complet, en définissant qu'un flot local  $(\phi^t)$  sur  $\Omega$  préserve une mesure  $\mu$  si pour tout  $t \in \mathbb{R}$  et tout ouvert  $U$  de  $\Omega$  sur lequel  $\phi^t$  est défini, nous avons  $(\phi^t)_*(\mu|_U) = \mu|_{\phi^t(U)}$ .

**Proposition 0.4. (Théorème de Liouville)** Soient  $m$  la mesure de Lebesgue sur  $\Omega$ , et  $\rho : \Omega \rightarrow [0, +\infty[$  une application de classe  $C^\infty$ . Le flot  $(\phi^t)_{t \in \mathbb{R}}$  préserve la mesure  $\rho dm$  si et seulement si (en notant  $F_k$  la  $k$ -ème coordonnée de  $F$ )

$$\sum_{k=1}^n \frac{\partial}{\partial x^k} (\rho F_k) = 0 .$$

**Démonstration.** Notons  $C_c^\infty(\Omega)$  l'ensemble des fonctions à valeurs réelles, de classe  $C^\infty$  et à support compact sur  $\Omega$ . Par la densité de  $C_c^\infty(\Omega)$  dans  $\mathbb{L}^1(\Omega, m)$ , il suffit de montrer que, pour tout  $t$  dans  $\mathbb{R}$  et tout  $f$  dans  $C_c^\infty(\Omega)$ , nous avons

$$\int_{\Omega} f \rho dm = \int_{\Omega} f \circ \phi^t \rho dm .$$

Comme  $\phi^0 = id$ , il suffit de montrer que, pour tout  $f$  dans  $C_c^\infty(\Omega)$ , nous avons

$$\frac{d}{dt} \left( \int_{\Omega} f \circ \phi^t \rho dm \right) = 0 .$$

Comme  $(\phi^t)_{t \in \mathbb{R}}$  est un groupe à un paramètre de difféomorphismes, il suffit de montrer que, pour tout  $f$  dans  $C_c^\infty(\Omega)$ , nous avons

$$\int_{\Omega} \frac{d}{dt} (f \circ \phi^t) |_{t=0} \rho dm = 0 .$$

Or, en notant  $I$  l'intégrale ci-dessus, nous avons par intégration par partie

$$I = \int_{\Omega} d_x f(F(x)) \rho(x) dm(x) = \int_{\Omega} \sum_{k=1}^n \rho F_k \frac{\partial f}{\partial x^k} dm = \int_{\Omega} \left( \sum_{k=1}^n \frac{\partial}{\partial x^k} (\rho F_k) \right) f dm .$$

Le résultat en découle. □

**Corollaire 0.5.** Le flot local d'un champ de vecteurs lisse sur un ouvert de  $\mathbb{R}^n$  préserve le volume si et seulement si sa divergence est nulle. □



33. Rudolf Lipschitz  
1832-1903



Hermann Schwarz  
1843-1921



Stefan Banach  
1892-1945



Frigyes Fiesz  
1880-1956

En mécanique classique, de nombreux systèmes physiques sont régis par un *système (d'équations différentielles) hamiltonien*, que nous allons définir. Soit  $\Omega$  un ouvert non vide de  $\mathbb{R}^{2n}$ , de coordonnées  $(q_1, \dots, q_n, p_1, \dots, p_n)$ , et  $H : \Omega \rightarrow \mathbb{R}$  une application de classe  $C^\infty$ , appelée un *hamiltonien*. Par exemple, dans le cas de l'attraction universelle,  $H$  est l'énergie totale, somme des énergies potentielle et cinétique,  $q = (q_1, \dots, q_n)$  est la position et  $p = (p_1, \dots, p_n)$  est la quantité de mouvement. Le système d'équations différentielles d'hamiltonien  $H$  est alors

$$\begin{cases} \forall i \in \llbracket 1, n \rrbracket, & \frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \\ \forall i \in \llbracket 1, n \rrbracket, & \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} . \end{cases}$$

Le champ de vecteurs  $F : (q = (q_1, \dots, q_n), p = (p_1, \dots, p_n)) \mapsto \left( \frac{\partial H}{\partial p_i}(p, q), -\frac{\partial H}{\partial q_i}(p, q) \right)$  d'un système hamiltonien est de divergence nulle, car par le lemme de Schwarz<sup>33</sup>, nous avons

$$\forall i \in \llbracket 1, n \rrbracket, \quad \frac{\partial}{\partial q_i} \left( \frac{\partial H}{\partial p_i} \right) + \frac{\partial}{\partial p_i} \left( -\frac{\partial H}{\partial q_i} \right) = 0 .$$

Donc par le corollaire 0.5, le flot d'un système hamiltonien complet sur  $\Omega$  préserve la mesure de Lebesgue sur  $\Omega$ .

L'hamiltonien est une *intégrale première* pour son flot, c'est-à-dire que les surfaces de niveau  $\Sigma_c$  définies par les équations  $H = c$ , pour  $c$  une constante, sont invariantes par le flot local  $(\phi^t)_t$  (au sens que pour tous les  $t \in \mathbb{R}$  et  $x \in \Sigma_c$  tels que  $\phi^t(x)$  est défini, nous avons  $\phi^t(x) \in \Sigma_c$ ). Lorsque  $\Sigma_c$  est une sous-variété compacte de l'ouvert  $\Omega$ , il est possible de montrer que le flot local en restriction à  $\Sigma_c$  est défini pour tout temps, et qu'il existe une mesure de probabilité naturelle sur  $\Sigma_c$ , qui est invariante par le flot hamiltonien.<sup>34</sup>

Un gaz idéal contenu dans une boîte parallélépipédique  $V$  est un exemple de tel système hamiltonien. Supposons que ce gaz suive les lois de la mécanique classique, deux molécules n'étant soumises qu'à une interaction répulsive  $U(r)$  fonction de leur distance mutuelle  $r$ . Supposons que les chocs sur les parois de la boîte soient élastiques. Si  $N$  est le nombre de molécules, l'espace des phases est une partie de  $\mathbb{R}^{6N}$  de coordonnées  $(p, q)$  avec  $p$  dans  $(\mathbb{R}^3)^N$  la quantité de mouvement et  $q$  dans  $(\mathbb{R}^3)^N$  la position des  $N$  particules. Ce gaz est donc un système hamiltonien pour l'hamiltonien

$$H = \sum_{i=1}^N \frac{p_i^2}{2m_i} + \sum_{1 \leq i \neq j \leq N} U(\|q_i - q_j\|) . \quad (7)$$

Si  $H_0$  est l'énergie initiale du système, l'évolution du gaz s'effectue sur l'hypersurface d'équation  $H = H_0$ , compacte car  $V$  l'est, et préserve donc une mesure de probabilité.

34. Nous renvoyons à un cours de géométrie différentielle et riemannienne, par exemple [Pau5] pour le contenu de cette note de bas de page. Cette mesure sur  $\Sigma_c$  est la mesure riemannienne pour la métrique riemannienne induite sur  $\Sigma_c$  par la métrique euclidienne de  $\mathbb{R}^n$ . Autrement dit, en considérant la  $n$ -forme différentielle  $dm = dx_1 \wedge dx_2 \wedge \dots \wedge dx_n$  sur  $\mathbb{R}^n$  (qui est la forme volume euclidienne), et  $X : \Sigma_c \rightarrow \mathbb{R}^n$  le champ de vecteurs unitaire (pour la norme euclidienne) orthogonal à  $\Sigma_c$  tel que  $d_x H(X(x)) > 0$  pour tout  $x \in \Sigma_c$ , alors la forme volume de l'hypersurface riemannienne transversalement orientée donc orientée  $\Sigma_c$  est  $i_X(dm)$ , et la mesure cherchée est la mesure définie par cette forme volume. L'invariance de cette mesure sur  $\Sigma_c$  découle (par un argument d'épaissement de  $\Sigma_c$ ) de l'invariance de la mesure de Lebesgue de  $\Omega$  par le flot local hamiltonien.

C'est une des motivations de la théorie ergodique d'arriver à comprendre le comportement statistique des orbites d'un système dynamique de la mécanique classique, précisément en fonction du fait qu'il laisse invariant une mesure.

### 0.3 Outils d'analyse fonctionnelle

Si  $(X, \phi)$  est un système dynamique topologique à temps discret, nous noterons

$$\text{Prob}_\phi(X)$$

l'ensemble des mesures de probabilité boréliennes sur  $X$  invariantes par  $\phi$ . Le résultat suivant garantit en particulier, sous des hypothèses de compacité, l'existence de mesure invariante.

Rappelons qu'un point  $x$  d'un convexe  $C$  d'un espace vectoriel réel  $E$  est dit *extrémal* dans  $C$  s'il n'existe aucun segment de  $E$  contenant  $x$  dans son intérieur et contenu dans  $C$ , c'est-à-dire si

$$\forall y, z \in C, \forall t \in ]0, 1[, \text{ si } x = ty + (1 - t)z \text{ alors } y = z .$$

Si  $X$  est un espace topologique compact métrisable, nous noterons  $\mathcal{M}_\mathbb{R}(X)$  l'espace vectoriel réel des mesures boréliennes signées<sup>35</sup> sur  $X$ , muni de la topologie vague<sup>36</sup>.

**Proposition 0.6.** *Soit  $X$  un espace topologique compact métrisable non vide, muni d'une application continue  $\phi : X \rightarrow X$ . Dans l'espace vectoriel topologique  $\mathcal{M}_\mathbb{R}(X)$  (pour la topologie vague), la partie  $\text{Prob}_\phi(X)$  est métrisable compacte convexe non vide<sup>37</sup>, et  $\text{Prob}_\phi(X)$  est l'enveloppe convexe fermée<sup>38</sup> de ses points extrémaux.*

---

35. Nous renvoyons par exemple à [Coh, §4] pour le contenu de cette note de bas de page. Une *mesure signée* sur un espace mesurable  $(X, \mathcal{A})$  est une application  $\mu : \mathcal{A} \rightarrow \mathbb{R}$ , telle que  $\mu(\emptyset) = 0$ , qui est  $\sigma$ -additive, c'est-à-dire telle que si  $(A_i)_{i \in I}$  est une famille dénombrable d'ensembles mesurables deux à deux disjoints, alors la série  $\sum_{i \in I} \mu(A_i)$  converge (c'est automatique si  $I$  est fini) et

$$\mu\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mu(A_i) .$$

Le théorème de décomposition de Jordan dit que toute mesure signée est la différence de deux mesures positives finies. La *variation*  $|\mu|$  d'une mesure signée  $\mu$  est la mesure positive finie sur  $(X, \mathcal{A})$  définie par

$$|\mu| : A \mapsto \sup_{(A_i)_{i \in I}} \sum_{i \in I} |\mu(A_i)| ,$$

où la borne supérieure est prise sur les partitions dénombrables  $(A_i)_{i \in I}$  de  $A$  par des éléments de  $\mathcal{A}$ .

36. c'est-à-dire, dans ce cas particulier où  $X$  est compact, de la topologie la moins fine rendant continues les applications

$$\mu \mapsto \mu(f) = \int_X f \, d\mu$$

pour toute application continue  $f \in C^0(X; \mathbb{R})$ . Ainsi  $\mathcal{M}_\mathbb{R}(X)$  est un *espace vectoriel (réel) topologique* (les applications de soustraction et de multiplication externe sont continues), et une suite  $(\mu_i)_{i \in \mathbb{N}}$  dans  $\mathcal{M}_\mathbb{R}(X)$  converge vers une mesure signée  $\mu \in \mathcal{M}_\mathbb{R}(X)$  si et seulement si la suite réelle  $(\mu_i(f))_{i \in \mathbb{N}}$  converge vers  $\mu(f)$  pour tout  $f \in C^0(X; \mathbb{R})$ . Lorsque la masse totale des mesures  $|\mu_i|$  est uniformément bornée, il suffit de vérifier ceci pour tout  $f$  dans une partie dense de  $C^0(X; \mathbb{R})$ .

37. Ce résultat d'existence d'au moins une mesure de probabilité invariante est parfois appelée le *théorème de Krylov-Bogolyubov*.

38. L'enveloppe convexe fermée d'une partie  $E$  d'un espace vectoriel topologique est l'intersection de tous les convexes fermés contenant  $E$ , c'est le plus petit convexe fermé contenant  $E$ .

**Démonstration.** Nous commençons par des rappels d'analyse fonctionnelle (voir par exemple [Coh] ou [Pau3, §6]), plus généraux que ce qui est strictement utile pour la démonstration de la proposition 0.6. Soit  $X$  un espace topologique localement compact.

Une application continue  $f : X \rightarrow \mathbb{R}$  s'annule à l'infini si pour tout  $\epsilon > 0$ , il existe un compact  $K$  de  $X$  tel que  $|f(x)| \leq \epsilon$  si  $x \notin K$ . Notons que si  $X$  est compact, alors toute application continue s'annule à l'infini. Notons  $C_0^0(X; \mathbb{R})$  l'espace de Banach<sup>33</sup> réel des applications continues de  $X$  dans  $\mathbb{R}$  qui s'annulent à l'infini, muni de la *norme uniforme*

$$f \mapsto \|f\|_\infty = \sup_{x \in X} |f(x)|,$$

et  $C_c^0(X; \mathbb{R})$  le sous-espace vectoriel des applications continues à support compact, qui est dense dans  $C_0^0(X; \mathbb{R})$ .

Notons  $\mathcal{M}_{\mathbb{R}}^r(X)$  l'espace de Banach des mesures boréliennes signées régulières<sup>39</sup> sur  $X$ , muni de la norme de la *variation totale*

$$\mu \mapsto \|\mu\| = |\mu|(X).$$

Par le *théorème de représentation de Riesz*<sup>33</sup> (voir [Coh, page 220]), l'application de  $\mathcal{M}_{\mathbb{R}}^r(X)$  dans le dual topologique<sup>40</sup>  $C_0^0(X; \mathbb{R})^\vee$  de  $C_0^0(X; \mathbb{R})$ , définie par

$$\mu \mapsto \{f \mapsto \mu(f) = \int_{x \in X} f(x) d\mu(x)\},$$

est un isomorphisme linéaire isométrique pour la norme de la variation totale sur  $\mathcal{M}_{\mathbb{R}}^r(X)$  et la norme duale<sup>41</sup> sur  $C_0^0(X; \mathbb{R})^\vee$ . Nous identifions désormais les deux espaces  $\mathcal{M}_{\mathbb{R}}^r(X)$  et  $C_0^0(X; \mathbb{R})^\vee$  par cette application.

La *topologie faible-étoile* sur l'espace vectoriel réel ou complexe  $C_0^0(X; \mathbb{R})^\vee$  est la topologie la moins fine rendant continues les applications  $\ell \mapsto \ell(f)$  d'évaluation en tout  $f \in C_0^0(X; \mathbb{R})$  des formes linéaires continues  $\ell$  sur  $C_0^0(X; \mathbb{R})$ . Elle fait de  $C_0^0(X; \mathbb{R})^\vee$  un espace vectoriel topologique localement convexe séparé.

La *topologie vague* sur l'espace vectoriel réel  $\mathcal{M}_{\mathbb{R}}^r(X)$  des mesures boréliennes signées régulières sur  $X$  est la topologie la moins fine rendant continues les applications

$$\mu \mapsto \mu(f) = \int_X f d\mu$$

pour toute application  $f \in \mathcal{C}_c^0(X; \mathbb{R})$  continue à support compact. Notons que si  $X$  est compact, alors toute application continue de  $X$  dans  $\mathbb{R}$  est à support compact. Puisque  $|\mu(f)| \leq \|\mu\| \|f\|_\infty$  pour tous les  $\mu \in \mathcal{M}_{\mathbb{R}}^r(X)$  et  $f \in \mathcal{C}_c^0(X; \mathbb{R})$ , la topologie vague est moins fine que la topologie induite par la norme de la variation totale (et donc toute suite dans  $\mathcal{M}_{\mathbb{R}}^r(X)$  qui converge en norme converge vaguement).

39. Nous ne rappelons pas ici la définition des mesures boréliennes régulières, nous mentionnons seulement que si  $X$  est localement compact et à base dénombrable d'ouverts (par exemple métrisable compact, voir la note de bas de page 79), alors toute mesure de probabilité est régulière, voir [Coh, page 206].

40. Le *dual topologique* d'un espace vectoriel réel ou complexe normé  $(E, \|\cdot\|)$  est l'espace vectoriel  $E^\vee$  des formes linéaires continues  $\ell$  sur  $E$ , muni de la *norme duale*

$$\|\ell\|^\vee = \sup_{x \in E, \|x\| \leq 1} |\ell(x)|.$$

41. Voir la note de bas de page précédente.

**Lemme 0.7.** *Étant donné une application continue propre<sup>42</sup>  $\varphi : Y \rightarrow Z$  entre deux espaces topologiques localement compacts, l'application  $\mu \mapsto \varphi_*\mu$  de mesure image par  $\varphi$ , de  $\mathcal{M}_{\mathbb{R}}^r(Y)$  dans  $\mathcal{M}_{\mathbb{R}}^r(Z)$ , est linéaire et continue pour la topologie vague.*

**Démonstration.** La linéarité est immédiate. Si  $f : Z \rightarrow \mathbb{R}$  est continue à support compact contenu dans  $K$ , alors  $f \circ \varphi : Y \rightarrow \mathbb{R}$  est continue à support contenu dans  $\varphi^{-1}(K)$ , qui est compact puisque  $\varphi$  est propre. Puisque  $\varphi_*\mu(f) = \mu(f \circ \varphi)$ , le résultat découle alors de la définition de la topologie vague.  $\square$

Une conséquence du *théorème de Banach-Alaoglu* de compacité de la boule unité fermée de  $C_0^0(X; \mathbb{R})$  pour la topologie faible-étoile donne le résultat suivant (voir par exemple [Pau3, §6] pour les compléments). Notons  $\text{Prob}(X)$  l'ensemble des mesures boréliennes régulières de probabilité sur  $X$ .

**Proposition 0.8.** *Les restrictions à  $\text{Prob}(X)$  des topologies vague et faible-étoile sur l'ensemble  $\mathcal{M}_{\mathbb{R}}^r(X)$  coïncident, et  $\text{Prob}(X)$  est convexe dans  $\mathcal{M}_{\mathbb{R}}^r(X)$ . Si  $X$  est compact, alors  $\text{Prob}(X)$  est compact pour la topologie vague (ou faible-étoile<sup>43</sup>). Si  $X$  est métrisable compact, alors la topologie vague sur  $\text{Prob}(X)$  est métrisable compacte.*  $\square$

Revenons maintenant à la démonstration de la proposition 0.6. Pour tous les  $t \in [0, 1]$  et  $\mu, \mu' \in \text{Prob}_{\phi}(X)$ , la mesure  $t\mu + (1-t)\mu'$  est clairement une mesure de probabilité invariante par  $\phi$ , par la linéarité de  $\phi_*$ . Donc  $\text{Prob}(X)$  est convexe. Par la continuité de  $\phi_*$  (voir le lemme 0.7), la partie  $\text{Prob}_{\phi}(X)$  est fermée dans le compact  $\text{Prob}(X)$ , donc compacte.

Montrons que  $\text{Prob}_{\phi}(X)$  est non vide. Partons d'une mesure de probabilité (forcément régulière)  $\nu$  sur  $X$  (par exemple la masse de Dirac<sup>44</sup>  $\Delta_x$  en un point  $x$  de l'espace  $X$ , qui est supposé non vide). Pour tout  $n \in \mathbb{N} - \{0\}$ , posons

$$\nu_n = \frac{1}{n} \sum_{k=0}^{n-1} (\phi^k)_*\nu,$$

qui est une mesure de probabilité. La compacité et la métrisabilité de  $\text{Prob}(X)$  (voir la proposition 0.8) assurent qu'il existe une sous-suite  $(\nu_{n_k})_{k \in \mathbb{N}}$  qui converge vers un élément  $\mu$  dans  $\text{Prob}(X)$ . Or la différence  $\nu_n - \phi_*(\nu_n) = \frac{1}{n}(\nu - (\phi^n)_*\nu)$  converge (pour la norme de la variation totale, car  $\nu_n - \phi_*(\nu_n)$  est de norme inférieure à  $\frac{2}{n}$ , donc) vaguement vers 0. Donc  $\phi_*\mu = \mu$  par passage à la limite et par le lemme 0.7 de continuité de  $\phi_*$ .

La dernière assertion de la proposition 0.6 découle du *théorème de Krein<sup>44</sup>-Milman* disant qu'un convexe compact d'un espace vectoriel topologique localement convexe séparé est l'enveloppe convexe fermée de ses points extrémaux.  $\square$

42. Une application continue entre deux espaces topologiques localement compacts est *propre* si l'image réciproque de tout compact est compact. Si l'espace de départ est compact, alors toute application continue est propre.

43. Ces deux topologies coïncident de manière évidente si  $X$  est compact, car alors nous avons  $C_c^0(X; \mathbb{R}) = C_0^0(X; \mathbb{R}) = C^0(X; \mathbb{R})$ .



44. Paul Dirac  
1902-1984



Mark Krein  
1907-1989



Henri Poincaré  
1854-1912



Marc Kac  
1914-1984



Eberhard Hopf  
1902-1983

## 0.4 Exercices

**Exercice E.2.** Soit  $X$  un espace topologique métrisable compact non vide, et considérons  $\phi_1, \phi_2 : X \rightarrow X$  deux applications continues qui commutent. Montrer qu'il existe une mesure de probabilité borélienne sur  $X$  invariante par  $\phi_1$  et  $\phi_2$ .

## 0.5 Indications pour la résolution des exercices

**Correction de l'exercice E.1.** Notons  $\Theta : \tilde{\Omega} \rightarrow \Omega^{\mathbb{N}}$  l'application  $\tilde{\omega} \mapsto (X_n(\tilde{\omega}))_{n \in \mathbb{N}}$ , qui est mesurable car chaque variable aléatoire  $X_n$  l'est, et vérifie que  $X_k = \text{pr}_k \circ \Theta$  pour tout  $k \in \mathbb{N}$ . Le fait que  $\Theta_*\mathbb{P} = \mu^{\mathbb{N}}$  traduit le fait que  $(X_n)_{n \in \mathbb{N}}$  est une suite de variables aléatoires indépendantes identiquement distribuées de loi  $\mu$ .

**Correction de l'exercice E.2.** Soit  $\nu$  une mesure de probabilité sur  $X$  (par exemple la masse de Dirac  $\Delta_x$  en un point  $x$  de  $X$ ). Pour tout  $n \in \mathbb{N}$ , posons

$$\nu_n = \frac{1}{(n+1)^2} \sum_{i,j=0}^n (\phi_1^i \phi_2^j)_* \nu = \frac{1}{(n+1)^2} \sum_{i,j=0}^n (\phi_2^i \phi_1^j)_* \nu,$$

qui est une mesure de probabilité. La compacité et métrisabilité de  $\text{Prob}(X)$  (voir la proposition 0.8) assure qu'il existe une sous-suite  $(\nu_{n_k})_{k \in \mathbb{N}}$  qui converge vers un élément  $\mu$  dans  $\text{Prob}(X)$ . Or la différence

$$\nu_n - (\phi_1)_*(\nu_n) = \frac{1}{(n+1)^2} \sum_{j=0}^n (\phi_2^j)_* (\nu - (\phi_1^{n+1})_* \nu),$$

dont la norme de la variation totale est inférieure à  $\frac{2(n+1)}{(n+1)^2}$ , converge vers 0 pour cette norme. Elle converge donc vaguement vers 0. Par conséquent,  $(\phi_1)_*\mu = \mu$  par passage à la limite et par le lemme 0.7. L'invariance de  $\mu$  par  $\phi_2$  se démontre de même.

# Première partie

## Théorie ergodique

La théorie ergodique est l'étude des systèmes dynamiques mesurables et mesurés, mais les outils topologiques, analytiques ou spectraux sont fort utiles (comme nous l'avons déjà vu dans la partie 0.2). Nous utiliserons principalement comme références les ouvrages [PY, Wal, CorFS, Pet, BenP, EW, Cou].

### 1 Généralités sur la dynamique mesurée

Nous donnons tout d'abord des définitions et des propriétés élémentaires des comportements presque sûrs des points et des événements pour les systèmes dynamiques probabilisés.

#### 1.1 Récurrence mesurée, errance mesurée, transience, conservativité

Soient  $(X, \mu, \phi)$  un système dynamique mesuré à temps discret et  $A$  une partie mesurable de  $X$ .

- Un point  $x \in X$  est dit (*infiniment positivement*) **récurrent dans**  $A$  si l'ensemble des  $n \in \mathbb{N}$  tels que  $\phi^n(x) \in A$  est infini.
- La partie mesurable  $A$  est (*presque partout infiniment positivement*) **récurrente** si  $\mu$ -presque tout point de  $A$  est (infiniment positivement) récurrent dans  $A$ . Par exemple, l'ensemble mesurable

$$R_A = A \cap \limsup_{n \in \mathbb{N}} \phi^{-n}(A)$$

des points de  $A$  qui reviennent une infinité de fois dans  $A$  est récurrent.

- La partie mesurable  $A$  est (*presque partout*) **errante**<sup>45</sup> si pour tout temps  $n \in \mathbb{N} \setminus \{0\}$ , nous avons  $\mu(\phi^{-n}(A) \cap A) = 0$ .<sup>46</sup> Par exemple, l'ensemble mesurable

$$E_A = A \setminus \bigcup_{n \in \mathbb{N}} \phi^{-n}(A)$$

des points de  $A$  qui ne reviennent jamais dans  $A$  est une partie errante, et si  $A$  est errante, alors  $\phi^{-n}(A)$  est aussi errante pour tout  $n \in \mathbb{N}$  (et pour tout  $n \in \mathbb{Z}$  si  $\phi$  est inversible).

- La partie mesurable  $A$  est (*presque partout*) **transiente** s'il existe une famille  $(E_i)_{i \in I}$ , indexée par un ensemble  $I$  dénombrable, d'ensembles errants tels que  $\mu(A \setminus \bigcup_{i \in I} E_i) = 0$ . Par exemple, l'ensemble mesurable

$$T_A = A \cap \bigcup_{n \in \mathbb{N}} \phi^{-n}(E_A)$$

des points de  $A$  qui à partir d'un certain temps ne reviennent pas dans  $A$  est une partie transiente (en posant  $I = \mathbb{N}$  et  $E_n = \phi^{-n}(E_A)$  pour tout  $n \in \mathbb{N}$ ). Comme tout point de  $A$

45. La lectrice ou le lecteur prendra bien garde à ne pas confondre cette notion d'errance mesurée avec la notion d'errance topologique qui sera introduite dans la partie II.

46. Ceci implique que  $\mu(\phi^{-n}(A) \cap \phi^{-m}(A)) = 0$  pour tous les  $n, m \in \mathbb{N}$  distincts.



ou bien revient une infinité de fois dans  $A$  ou bien à partir d'un certain temps n'y revient pas, nous avons une partition

$$A = R_A \sqcup T_A = \left( A \cap \limsup_{n \in \mathbb{N}} \phi^{-n}(A) \right) \sqcup \left( A \cap \bigcup_{n \in \mathbb{N}} \phi^{-n}(E_A) \right) \quad (8)$$

de  $A$  entre une partie mesurable de  $A$  qui est récurrente et une partie mesurable de  $A$  qui est transiente.

• Nous dirons que le système dynamique mesuré  $(X, \mu, \phi)$  est *conservatif* si tout ensemble errant est de mesure nulle, et *complètement dissipatif* si l'ensemble  $X$  tout entier est transient. Par exemple, si  $\mu$  est une mesure finie (par exemple de probabilité), alors  $(X, \mu, \phi)$  est conservatif. En effet, si  $E$  est un ensemble errant de mesure non nulle, alors, par la  $\sigma$ -additivité et la  $\phi$ -invariance, nous avons

$$\mu(X) \geq \mu\left(\bigcup_{n \in \mathbb{N}} \phi^{-n}(E)\right) = \sum_{n \in \mathbb{N}} \mu(\phi^{-n}(E)) = \sum_{n \in \mathbb{N}} \mu(E) = \infty.$$

Par exemple, la translation entière  $\lambda_1 : t \mapsto t + 1$  sur  $\mathbb{R}$  est complètement dissipative pour la mesure de Lebesgue, car  $[0, 1]$  est alors une partie errante et  $X = \bigcup_{n \in \mathbb{Z}} \lambda_1^{-n}([0, 1])$ .

### **Théorème 1.1. (Théorème de récurrence de Poincaré <sup>44</sup>)**

- (1) Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret, et  $A$  un élément de  $\mathcal{B}$ . Alors, pour  $\mu$ -presque tout point  $x$  de  $A$ , l'orbite  $(\phi^n(x))_{n \in \mathbb{N}}$  repasse une infinité de fois dans  $A$ .
- (2) Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique mesuré à temps discret conservatif, et  $A$  un élément de  $\mathcal{B}$ . Alors, pour  $\mu$ -presque tout point  $x$  de  $A$ , l'orbite  $(\phi^n(x))_{n \in \mathbb{N}}$  repasse une infinité de fois dans  $A$ .

Bien entendu, l'assertion (1) est un cas particulier de l'assertion (2). Mais donner deux arguments est pédagogique.

**Démonstration.** (1) Notons  $A_n = \bigcup_{k \geq n} \phi^{-k}(A)$  l'ensemble des points de  $X$  qui passent dans  $A$  lorsque nous itérons  $\phi$  au moins  $n$  fois. Notons  $A_\infty = \bigcap_{n \in \mathbb{N}} A_n$  l'ensemble des points qui passent une infinité de fois dans  $A$ . Alors  $\mu(A_{n+1}) = \mu(\phi^{-1}(A_n)) = \mu(A_n)$ , et  $A \subset A_0$ . Puisque la suite  $(A_n)_{n \in \mathbb{N}}$  est décroissante et comme  $\mu(A_n - A_0) = 0$ , nous avons  $\mu(A_n \cap A) = \mu(A_0 \cap A) = \mu(A)$ . Par conséquent  $\mu(A_\infty \cap A) = \mu(A)$  par les propriétés des mesures finies. Ceci démontre le résultat.

(2) Puisque l'ensemble  $E_A$  est errant, il est de mesure nulle. Donc  $T_A$ , qui est une union dénombrable d'ensembles de mesure nulle, est de mesure nulle. D'où  $\mu(A \Delta R_A) = 0$  par la formule (8), ce qui donne le résultat.  $\square$

## **1.2 Temps de retour et théorème de Kac <sup>44</sup>**

Le théorème de Poincaré ci-dessus fournit des points qui reviennent infiniment souvent dans une partie mesurable donnée. Mais il ne donne aucun contrôle sur le temps passé entre deux retours successifs. Dans cette partie, nous allons, après quelques définitions, étudier ce problème en moyenne.

Soient  $(X, \mu, \phi)$  un système dynamique mesuré à temps discret et  $A$  une partie mesurable de  $X$  telle que  $0 < \mu(A) < +\infty$ .

Le *premier temps de retour* de  $\phi$  dans  $A$  est l'application mesurable  $\tau_A$  de  $X$  dans l'ensemble discret  $(\mathbb{N} \setminus \{0\}) \cup \{+\infty\}$  définie pour tout  $x \in X$  par

$$\tau_A(x) = \min\{n \in \mathbb{N} \setminus \{0\} : \phi^n(x) \in A\}$$

avec la convention usuelle que le minimum d'une partie vide de  $\mathbb{N}$  est  $+\infty$ . En notant  $\mu_A$  la restriction de  $\mu$  à  $A$  normalisée pour être de probabilité<sup>47</sup>, la *loi du premier temps de retour dans  $A$*  est la mesure sur  $(\mathbb{N} \setminus \{0\}) \cup \{+\infty\}$  qui est la mesure image  $(\tau_A)_*(\mu_A)$  de  $\mu_A$  par  $\tau_A$  : pour tout  $n \in (\mathbb{N} \setminus \{0\}) \cup \{+\infty\}$ , nous avons

$$(\tau_A)_*(\mu_A)(\{n\}) = \frac{\mu(A \cap \tau_A^{-1}(n))}{\mu(A)}.$$

De manière similaire, posons par récurrence  $\tau_A^1 = \tau_A$  et pour tous les  $k \geq 1$  et  $x \in X$ ,

$$\tau_A^{k+1}(x) = \min\{n > \tau_A^k(x) : \phi^n(x) \in A\}$$

(avec la même convention). L'application mesurable  $\tau_A^k : X \rightarrow (\mathbb{N} \setminus \{0\}) \cup \{+\infty\}$  ainsi définie est appelée le  *$k$ -ème temps de retour* dans  $A$ . Comme le  $(k+1)$ -ème retour d'un point est le premier retour du  $k$ -ème retour de ce point, nous avons

$$\forall x \in X, \quad \tau_A^{k+1}(x) = \tau_A^k(x) + \tau_A(\phi^{\tau_A^k(x)}(x)). \quad (9)$$

Si la mesure  $\mu$  est finie, alors le théorème de récurrence de Poincaré 1.1 dit que pour tout  $k \in (\mathbb{N} \setminus \{0\}) \cup \{+\infty\}$  et pour  $\mu$ -presque tout  $x \in A$ , le  $k$ -ème temps de retour  $\tau_A^k(x)$  de  $x$  dans  $A$  est fini.

Le résultat suivant permet de comprendre quel est en moyenne le premier temps de retour.

**Théorème 1.2. (Théorème de Kac)** *Si  $\mu$  est une mesure de probabilité, alors*

$$\frac{1}{\mu(A)} \int_A \tau_A d\mu = \frac{\mu(\tau_A^{-1}(\mathbb{N} \setminus \{0\}))}{\mu(A)}.$$

En particulier, si  $\mu$ -presque tout point de  $X$  est récurrent dans  $A$ , alors l'espérance conditionnelle au fait de partir de  $A$  du temps de premier retour dans  $A$  est égale à  $\frac{1}{\mu(A)}$ .

**Démonstration.** Pour tout entier  $n \geq 1$ , notons  $A_n = A \cap \tau_A^{-1}(n)$  et  $X_n = {}^c A \cap \tau_A^{-1}(n)$  le complémentaire de  $A_n$  dans  $\tau_A^{-1}(n)$ . En utilisant dans le calcul ci-dessous

- l'additivité de  $\mu$ , pour la première égalité,
- l'invariance de  $\mu$  par  $\phi$ , pour la deuxième égalité,
- le fait que  $\phi^{-1}(X_n) = \tau_A^{-1}(n+1)$  puisque  $\phi^{-1}(X_n)$  est l'ensemble des éléments de  $X$  dont l'image par  $\phi$  d'une part n'appartient pas à  $A$ , d'autre part appartient à  $\tau_A^{-1}(n)$ , pour la troisième égalité,
- une itération et puisque  $\lim_{n \rightarrow +\infty} \mu(\tau_A^{-1}(n)) = 0$  car les parties mesurables  $\tau_A^{-1}(n)$  sont deux à deux disjointes et la mesure est finie, pour la quatrième égalité,
- la  $\sigma$ -additivité de  $\mu$ , pour la cinquième égalité,
- le fait que  $\mu$ -presque tout point  $x \in A$  est récurrent dans  $A$  par le théorème de récurrence de Poincaré 1.1 (1), pour la dernière égalité,

---

47. qui est la mesure conditionnelle à  $A$  de  $\mu$  lorsque  $\mu$  est une mesure de probabilité

nous avons

$$\begin{aligned} \mu(\tau_A^{-1}(n)) &= \mu(A_n) + \mu(X_n) = \mu(A_n) + \mu(\phi^{-1}(X_n)) = \mu(A_n) + \mu(\tau_A^{-1}(n+1)) \\ &= \sum_{k \geq n} \mu(A_k) = \mu\left(\bigcup_{k \geq n} A_k\right) = \mu(\{x \in A : n \leq \tau_A(x) < \infty\}) = \mu(\{x \in A : \tau_A(x) \geq n\}). \end{aligned}$$

D'où par la  $\sigma$ -additivité de  $\mu$  pour la première égalité et par la définition de l'intégrale des fonctions étagées pour la dernière égalité, nous avons

$$\mu(\tau_A^{-1}(\mathbb{N} \setminus \{0\})) = \sum_{n=1}^{+\infty} \mu(\tau_A^{-1}(n)) = \sum_{n=1}^{+\infty} \mu(\{x \in A : \tau_A(x) \geq n\}) = \int_A \tau_A d\mu.$$

Le résultat en découle en divisant par  $\mu(A)$ .  $\square$

### 1.3 Exercices

**Exercice E.3.** Soit  $(X, \mu, \phi)$  un système dynamique mesuré à temps discret, où  $\mu$  est  $\sigma$ -finie. Montrer qu'il existe un couple<sup>48</sup>  $(C, D)$  de parties de  $X$  tel que

- (1) les parties  $C$  et  $D$  sont mesurables, invariantes par  $\phi$  et  $X = C \sqcup D$  est la réunion disjointe de  $C$  et  $D$ ,
- (2) le système dynamique restreint  $(C, \mu|_C, \phi|_C)$  est conservatif;
- (3) le système dynamique restreint  $(D, \mu|_D, \phi|_D)$  est totalement dissipatif;
- (4) Si  $(C', D')$  est un autre couple de parties de  $X$  vérifiant les propriétés (1), (2) et (3), alors

$$\mu(C \Delta C') = \mu(D \Delta D') = 0.$$

**Exercice E.4.** Soient  $\mathcal{A}$  un ensemble fini discret de cardinal au moins 2, et  $\mu$  une mesure de probabilité sur  $\mathcal{A}$ . Soit  $a \in \mathcal{A}$ . Calculer les lois du premier temps de retour sur un cylindre  $[a]$  des décalages de Bernoulli bilatère et unilatère d'alphabet  $\mathcal{A}$  pour les mesures de Bernoulli  $\mu^{\mathbb{Z}}$  et  $\mu^{\mathbb{N}}$  respectivement. Vérifier le théorème de Kac 1.2.

**Exercice E.5.** Soient  $\alpha \in \mathbb{R}/\mathbb{Z}$  et  $\phi$  la translation du cercle  $\mathbb{R}/\mathbb{Z}$  définie par  $x \mapsto x + \alpha$ . Soient  $\beta \in ]0, 1[$  et  $I$  l'intervalle  $[0, \beta]$  de  $\mathbb{R}$  identifié avec son image dans  $\mathbb{R}/\mathbb{Z}$ .

- (1) Montrer que pour tous les  $a < b$  dans  $I$ ,  $x \in [a, b]$  et  $n \in \mathbb{N}$ , si  $\phi^n(x) \in I$ , alors  $\phi^n(a) \in I$  ou  $\phi^n(b) \in I$ .
- (2) Montrer que le temps de retour de  $\phi$  sur  $I$  prend au plus trois valeurs.
- (3) Déterminer quand le temps de retour de  $\phi$  sur  $I$  ne prend que deux valeurs distinctes.

### 1.4 Indications pour la résolution des exercices

**Correction de l'exercice E.3.** Voir par exemple [Aar, §1.1].

**Correction de l'exercice E.4.** Nous traitons le cas unilatère, le cas bilatère est similaire. Par l'invariance par décalage, nous pouvons supposer que  $[a] = \{(x_i)_{i \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}} : x_0 = a\}$ . Alors par la définition du premier temps de retour, pour tout  $n \in \mathbb{N} \setminus \{0\}$ , nous avons

$$\tau_{[a]}^{-1}(n) = \{(x_i)_{i \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}} : x_1 \neq a, \dots, x_{n-1} \neq a, x_n = a\}.$$

48. Un tel couple est appelé une *décomposition de Hopf*<sup>44</sup> de  $(X, \mu, \phi)$ , et la propriété (4) sa propriété d'unicité (modulo zéro, ou modulo ensembles de mesure nulle).

Par la définition de la mesure de Bernoulli  $\mu^{\mathbb{N}}$ , nous avons donc

$$\begin{aligned} (\tau_{[a]}^*) (\mu_{[a]}^{\mathbb{N}}) (n) &= \frac{1}{\mu^{\mathbb{N}}([a])} \mu^{\mathbb{N}}([a] \cap (\tau_{[a]}^{-1}(n))) = \frac{1}{\mu(a)} \left( \mu(a) \prod_{i=1}^{n-1} (1 - \mu(a)) \mu(a) \right) \\ &= (1 - \mu(a))^{n-1} \mu(a) . \end{aligned}$$

(avec la convention usuelle qu'un produit vide vaut 1). Il s'agit donc de la loi géométrique de paramètre  $\mu(a)$  sur  $\mathbb{N} \setminus \{0\}$ .

Nous avons

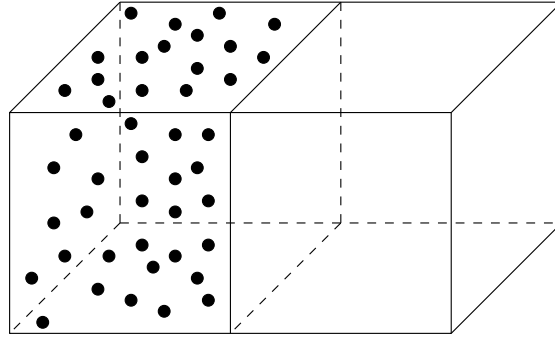
$$\begin{aligned} \frac{1}{\mu^{\mathbb{N}}([a])} \int_{[a]} \tau_A d\mu^{\mathbb{N}} &= \frac{1}{\mu^{\mathbb{N}}([a])} \sum_{n=1}^{\infty} n \mu^{\mathbb{N}}([a] \cap \tau_{[a]}^{-1}(n)) = \sum_{n=1}^{\infty} n (1 - \mu(a))^{n-1} \mu(a) \\ &= \frac{1}{\mu(a)} = \frac{1}{\mu^{\mathbb{N}}([a])} . \end{aligned}$$

conformément au théorème de Kac 1.2 (lorsque la transformation est ergodique).

## 2 Ergodicité

Historiquement, le mot ergodique (en grec  $\epsilon\rho\gamma\omicron\nu$  = énergie,  $\omicron\delta\acute{o}\varsigma$  = chemin) reflète l'idée, due à Boltzmann, que, le plus souvent, un système physique décrit au cours du temps un chemin sur l'hypersurface à énergie constante, chemin qui est dense sur cette hypersurface lorsque le système est assez chaotique. L'exercice E.12 renforce l'intuition "je passe partout" sous-jacente au mot ergodique. De nos jours, le concept mathématique d'ergodicité ne fait plus référence à la physique ou à la topologie.

Donnons toutefois une illustration historique en physique statistique.



Considérons l'évolution d'une mole de gaz parfait (donc avec environ  $N = 6 \cdot 10^{23}$  particules) dans une boîte (avec chocs élastique sur les bords), qui au temps  $t = 0$  est dans la moitié gauche de la boîte. Notons  $A$  l'événement de l'espace des phases où toutes les particules de gaz sont dans la moitié gauche, qui comme chaque particule est ou bien à gauche ou bien à droite, est de volume de l'ordre de  $2^{-N} \approx 10^{-18}$ . Par le corollaire 0.5 du théorème de Liouville de préservation, par le système hamiltonien d'hamiltonien donné par la formule (7), du volume normalisé  $\mu$ , et par le théorème de récurrence de Poincaré 1.1, nous nous attendons donc à ce que le gaz revienne entièrement dans la moitié gauche de la boîte. Mais au bout de combien de temps ?

Par l'hypothèse ergodique de Boltzman (revisitée par Tatiana Afanassieva et Paul Ehrenfest) en physique statistique (dont nous verrons dans l'exercice E.6 qu'elle implique que nous avons  $\mu(\{\tau_A < +\infty\}) = 1$ ), et par le théorème de Kac 1.2, le temps moyen espéré est en fait de l'ordre de  $2^6 \cdot 10^{36}$ , soit probablement plus long que la durée de vie de l'univers !

## 2.1 Transformations ergodiques

Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique mesuré. Un exemple d'obstruction à la propriété physique ci-dessus que presque toute orbite "passe partout" est celui où l'espace des phases est réunion disjointes de deux parties invariantes de mesure non nulle.

Nous dirons que le système dynamique  $(X, \mathcal{B}, \mu, \phi)$  est *ergodique*, ou que la transformation (en temps discret) ou le flot (en temps continu)  $\phi$  de l'espace mesuré  $(X, \mathcal{B}, \mu)$  est ergodique, ou que la mesure  $\mu$  sur  $(X, \mathcal{B})$  est ergodique pour  $\phi$ , si

$$\forall A \in \mathcal{B}, \quad \phi^{-1}(A) = A \Rightarrow \mu(A) = 0 \text{ ou } \mu({}^c A) = 0$$

en temps discret et

$$\forall A \in \mathcal{B}, \quad (\forall t \in \mathbb{R}, \quad \phi^{-t}(A) = A) \Rightarrow \mu(A) = 0 \text{ ou } \mu({}^c A) = 0$$

en temps continu. Lorsque  $\mu$  est une mesure de probabilité, cette dernière condition s'écrit  $\mu(A) = 0$  ou  $1$ . Lorsque la transformation ou le flot  $\phi$  est sous-entendu, c'est la mesure  $\mu$  que nous qualifions d'*ergodique*. Voir [Aar] pour l'extension au cas où  $\phi$  ne préserve pas  $\mu$ , mais préserve les ensembles de mesure nulle.

La proposition suivante donne des propriétés équivalentes pour l'ergodicité du système dynamique  $(X, \mathcal{B}, \mu, \phi)$ . Elle caractérise en particulier les points extrémaux du convexe compact non vide  $\text{Prob}_\phi(X)$  défini dans la partie 0.2.

**Proposition 2.1.** *Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé, et  $p \in [1, +\infty[$ .<sup>49</sup> Les quatre assertions suivantes sont équivalentes.*

- (1) *Le système dynamique  $(X, \mathcal{B}, \mu, \phi)$  est ergodique.*
- (2) *Toute application  $f : X \rightarrow \mathbb{C}$  mesurable, telle que  $f \circ \phi = f$  presque partout, est presque partout constante.*
- (3) *Toute application  $f$  dans  $\mathbb{L}^1(X, \mu)$ , telle que  $f \circ \phi = f$  presque partout, est presque partout constante.*
- (4) *Toute application  $f$  dans  $\mathbb{L}^p(X, \mu)$ , telle que  $f \circ \phi = f$  presque partout, est presque partout constante.*

*Lorsque  $X$  est un espace topologique métrisable compact, non vide, ces assertions sont équivalentes à la suivante.*

- (5) *La mesure  $\mu$  est un point extrémal du convexe  $\text{Prob}_\phi(X)$  des mesures de probabilité invariantes par  $\phi$  sur  $(X)$ , dans l'espace vectoriel  $\mathcal{M}_\mathbb{R}(X)$  des mesures signées sur  $(X)$ .*

Lorsque le système dynamique probabilisé  $(X, \mathcal{B}, \mu, \phi)$  est à temps continu, rappelons que l'invariance  $f \circ \phi = f$  presque partout signifie que pour tout  $t \in \mathbb{R}$ , nous avons  $f \circ \phi^t = f$  presque partout. Nous ne donnons la démonstration que dans le cas du temps discret, mais celle en temps continu est analogue.

49. C'est surtout le cas  $p = 2$  qui est utile en pratique, car alors  $\mathbb{L}^p(X, \mu)$  est un espace de Hilbert<sup>50</sup>.

L'assertion (4) donne en particulier une interprétation  $\mathbb{L}^2$  de l'ergodicité d'une transformation  $\phi$ , voir la partie 3.3 pour des développements.

**Démonstration.** Il est immédiat, car  $\mathbb{L}^p(X, \mu) \subset \mathbb{L}^1(X, \mu)$  (car  $\mu$  est finie et par l'inégalité de Hölder<sup>50</sup>) et en prenant pour  $f \in \mathbb{L}^p(X, \mu)$  la fonction indicatrice<sup>51</sup> (ou application caractéristique)  $\mathbb{1}_A$  de  $A$ , que (2) implique (3) implique (4) implique (1).

Pour montrer que (1) implique (2), nous pouvons supposer que  $f$  est à valeurs réelles. Si  $f$  n'est pas constante presque partout, il existe un réel  $x$  tel que, en notant  $A' = f^{-1}([x, +\infty[)$ , nous avons  $0 < \mu(A') < 1$ . L'égalité  $f \circ \phi = f$  presque partout assure que les ensembles  $A'$  et  $\phi^{-1}(A')$  coïncident en dehors d'une partie négligeable. L'ensemble mesurable  $A = \bigcap_{p \in \mathbb{N}} \bigcup_{n \geq p} \phi^{-n}(A')$  coïncide alors avec  $A'$  en dehors d'une partie négligeable. En particulier, nous avons  $0 < \mu(A) < 1$ . Comme  $\phi^{-1}(A) = A$ , ceci contredit l'ergodicité de  $\mu$ .

Pour montrer que les assertions (1) et (5) sont équivalentes, montrons que  $\mu$  est ergodique si et seulement si, pour toutes les mesures de probabilité  $\mu_1, \mu_2$  invariantes par  $\phi$  sur  $X$  et tout  $t$  dans  $]0, 1[$  tel que  $\mu = t\mu_1 + (1-t)\mu_2$ , nous avons  $\mu = \mu_1 = \mu_2$ .

Supposons d'abord  $\mu$  ergodique. Soient  $\mu_1, \mu_2 \in \text{Prob}_\phi(X)$  et  $t \in ]0, 1[$  tels que  $\mu = t\mu_1 + (1-t)\mu_2$ . Les ensembles de mesure nulle pour  $\mu$  sont aussi de mesure nulle pour  $\mu_1$ . Le théorème<sup>52</sup> de Radon<sup>50</sup>-Nikodym<sup>50</sup> (voir [Coh, page 132]) assure qu'il existe une fonction  $f$  dans  $\mathbb{L}^1(X, \mu)$  telle que  $\mu_1 = f\mu$ . Comme  $\mu$  et  $\mu_1$  sont invariantes par  $\phi$ , nous avons  $f = f \circ \phi$  presque partout pour  $\mu$ . Comme  $\mu$  est ergodique, l'application  $f$  est constante presque partout et les mesures de probabilité  $\mu$  et  $\mu_1$  sont égales. Donc  $\mu_1 = \mu_2$  et  $\mu$  est extrémale.

Réciproquement, supposons que  $\mu$  ne soit pas ergodique. Soit  $A_1$  une partie mesurable de  $X$  telle que nous ayons  $\phi^{-1}(A_1) = A_1$  et  $0 < \mu(A_1) < 1$ , et soit  $A_2 = X \setminus A_1$ . Posons alors  $\mu_i = \frac{1}{\mu(A_i)} \mu|_{A_i}$  pour  $i = 1, 2$ . Les mesures de probabilité  $\mu_1$  et  $\mu_2$  sont invariantes par  $\phi$ , différentes de  $\mu$  et  $\mu = \mu(A_1)\mu_1 + (1 - \mu(A_1))\mu_2$ . Donc  $\mu$  n'est pas extrémale.  $\square$

**Exercice E.6.** Montrer que si  $(X, \mathcal{B}, \mu, \phi)$  est un système dynamique mesuré à temps discret conservatif et ergodique, alors pour toute partie mesurable  $B$  de mesure non nulle, presque tout point de  $X$  est (positivement infiniment) récurrent dans  $B$ .

## 2.2 Le théorème ergodique de Birkhoff<sup>50</sup>

La propriété d'ergodicité d'un système dynamique mesuré est souvent utilisée en conjonction avec le théorème suivant, qui est une version « quantitative » du théorème de récurrence de Poincaré 1.1, celui-ci impliquant, sous l'hypothèse d'ergodicité (voir l'exercice (E.6) ci-dessus) que presque toute orbite passe infiniment souvent dans toutes partie de



David Hilbert  
50. 1842-1943



Otto Hölder  
1859-1937



Johan Radon  
1887-1956



Otto Nikodym  
1887-1974



Georges Birkhoff  
1884-1944

51. Rappelons que  $\mathbb{1}_A(x)$  vaut 1 si  $x$  appartient à  $A$  et 0 sinon.

52. qui s'applique car  $\mu_1$  et  $\mu_2$  sont des mesures de probabilité, donc des mesures  $\sigma$ -finies

mesure non nulle. Il affirme en particulier que, lorsque le système est ergodique, pour toute partie mesurable  $A$ , la proportion de temps passé dans  $A$  par presque toutes les orbites est égale à  $\mu(A)$  (voir la remarque (ii) ci-dessous).

**Théorème 2.2. (Théorème ergodique de Birkhoff)** *Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique mesuré. Pour tout  $f$  dans  $\mathbb{L}^1(X, \mu)$ , pour tout  $T \in \mathbb{N} - \{0\}$  si le système dynamique est à temps discret (respectivement pour tout  $T > 0$  si le système dynamique est à temps continu), notons*

$$S_T f(x) = \frac{1}{T} \sum_{k=0}^{T-1} f(\phi^k(x))$$

(respectivement

$$S_T f(x) = \frac{1}{T} \int_0^T f(\phi^t(x)) dt).$$

(1) *La limite  $\tilde{f}(x) = \lim_{T \rightarrow \infty} S_T f(x)$  existe pour  $\mu$ -presque tout  $x \in X$ .*

(2)  *$\tilde{f} \circ \phi = \tilde{f}$  presque partout.*

(3)  *$\|\tilde{f}\|_{\mathbb{L}^1} \leq \|f\|_{\mathbb{L}^1}$ , et en particulier l'application linéaire  $f \mapsto \tilde{f}$  de  $\mathbb{L}^1(X, \mu)$  dans  $\mathbb{L}^1(X, \mu)$  est continue.*

(4) *Si la mesure  $\mu$  est finie, alors la convergence a lieu dans  $\mathbb{L}^1$  :*

$$\lim_{T \rightarrow \infty} \|S_T f - \tilde{f}\|_{\mathbb{L}^1} = 0.$$

(5) *Pour toute partie  $A$  dans  $\mathcal{B}$  invariante par  $\phi$ , de mesure finie, nous avons*

$$\int_A f d\mu = \int_A \tilde{f} d\mu.$$

(6) *Si  $\mu$  est une mesure de probabilité ergodique, alors*

$$\tilde{f}(x) = \lim_{T \rightarrow \infty} S_T f(x) = \int_X f d\mu$$

*pour  $\mu$ -presque tout  $x \in X$ , et en particulier  $\tilde{f}$  est presque partout constante.*

Donnons quelques commentaires avant de proposer une démonstration du théorème ergodique de Birkhoff.

**Remarques.** (i) La quantité  $S_T f(x)$  s'appelle une *moyenne de Birkhoff* de  $f$ .<sup>53</sup> La limite (lorsqu'elle existe)  $\lim_{T \rightarrow \infty} S_T f(x)$  s'appelle la *moyenne orbitale* (ou *temporelle*). L'intégrale  $\int_X f d\mu$  s'appelle la *moyenne spatiale* de  $f$ . L'assertion (6) du théorème ergodique de Birkhoff dit donc, en paraphrasant, que si  $\mu$  est une probabilité ergodique, alors presque toutes les moyennes temporelles d'une fonction intégrable coïncident avec sa moyenne spatiale. Ceci traduit une propriété « d'équirépartition » de presque toute orbite (voir

<sup>53</sup>. Certains ouvrages notent plutôt  $S_T f(x)$  la somme ou intégrale non renormalisée  $\sum_{k=0}^{T-1} f(\phi^k(x))$  ou  $\int_0^T f(\phi^t(x)) dt$ .

aussi<sup>54</sup> la partie 4.2), au sens que nous pouvons calculer les intégrales des fonctions en prenant des moyennes des valeurs de cette fonction sur les orbites.

(ii) Si  $\mu$  est une mesure de probabilité ergodique, en prenant pour  $f$  l'application caractéristique  $\mathbb{1}_A$  d'une partie mesurable  $A$  de  $X$ , l'assertion (6) du résultat précédent dit que la proportion de temps

$$\lim_{T \rightarrow \infty} \frac{1}{T} \text{Card}\{k \leq T : \phi^n(x) \in A\} \quad (\text{respectivement } \lim_{T \rightarrow \infty} \frac{1}{T} \text{Leb}\{t \leq T : \phi^t(x) \in A\})$$

que l'orbite de  $x$  passe dans  $A$  est égale à  $\mu(A)$ , pour  $\mu$ -presque tout point  $x$  de  $X$ .

(iii) Notons  $\mathcal{I}$  la *tribu des invariants* du système dynamique mesuré  $(X, \mathcal{B}, \mu, \phi)$ , c'est-à-dire la tribu engendrée par l'ensemble des éléments de  $\mathcal{B}$  presque invariants par  $\phi$ , ou encore la plus petite sous-tribu  $\mathcal{I}$  de  $\mathcal{B}$  rendant  $\mathcal{I}$ -mesurables les applications  $\mathcal{B}$ -mesurables presque partout invariables. En particulier, si  $A \in \mathcal{B}$  est une partie presque  $\phi$ -invariante, alors  $A \in \mathcal{I}$ . L'application  $\mathcal{B}$ -mesurable  $\mu$ -presque partout invariante  $\tilde{f}$  donnée par l'assertion (1) est en fait égale à l'espérance conditionnelle<sup>55</sup>  $\mathbb{E}(f \mid \mathcal{I})$   $\mu$ -presque partout. En effet, l'application  $\tilde{f}$  est  $\mathcal{I}$ -mesurable par l'assertion (2). De plus, pour tout  $B \in \mathcal{I}$ , par l'invariance de  $\mu$ , nous avons  $\mathbb{E}[f \circ \phi \mathbb{1}_B] = \mathbb{E}[f \mathbb{1}_B]$ . Donc par la linéarité de l'espérance et par le théorème de convergence dominée de Lebesgue lorsque  $f$  est bornée, et par densité sinon (voir la démonstration de l'assertion (4)), nous avons

$$\forall B \in \mathcal{I}, \quad \mathbb{E}[\mathbb{E}[f \mid \mathcal{I}] \mathbb{1}_B] = \mathbb{E}[f \mathbb{1}_B] = \mathbb{E}[S_n f \mathbb{1}_B] \rightarrow_{n \rightarrow \infty} \mathbb{E}[\tilde{f} \mathbb{1}_B].$$

Le résultat découle alors de la propriété d'unicité de l'espérance conditionnelle.

**Démonstration du théorème de Birkhoff 2.2.** Nous ne la donnons que dans le cas du temps discret, celle en temps continu est analogue. Pour  $f$  dans  $\mathbb{L}^1(X, \mu)$ , notons

$$f^*(x) = \sup_{n \geq 1} \frac{1}{n} \sum_{k=0}^{n-1} f(\phi^k(x)).$$

**Lemme 2.3. (Lemme ergodique maximal)** *Pour tout  $f$  dans  $\mathbb{L}^1(X, \mu)$ , nous avons*

$$\int_{\{x \in X : f^*(x) > 0\}} f \, d\mu \geq 0.$$

---

54. en particulier pour des situations où il est possible d'éviter la restriction à seulement presque toute orbite

55. Soient  $(\Omega, \mathcal{A}, \mu)$  un espace de probabilité,  $X \in \mathbb{L}^1(\Omega, \mathcal{A}, \mu)$  une variable aléatoire  $\mathcal{A}$ -mesurable  $\mu$ -intégrable, et  $\mathcal{B}$  une sous-tribu de  $\mathcal{A}$ . Rappelons (voir par exemple [LeG, §11.2][Neve]) que l'espérance conditionnelle de  $X$  sachant  $\mathcal{B}$  est l'unique (modulo égalité  $\mu$ -presque partout) variable aléatoire  $\mathcal{B}$ -mesurable  $\mu$ -intégrable  $\mathbb{E}[X \mid \mathcal{B}] \in \mathbb{L}^1(\Omega, \mathcal{B}, \mu)$  telle que

$$\forall B \in \mathcal{B}, \quad \mathbb{E}[X \mathbb{1}_B] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{B}] \mathbb{1}_B]$$

ou, de manière équivalente, pour toute variable aléatoire réelle  $\mathcal{B}$ -mesurable bornée  $Y$ ,

$$\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{B}] Y].$$

En particulier,  $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{B}]]$ . L'espérance conditionnelle est  $\mu$ -presque partout positive ou nulle si  $X$  l'est. Si  $X$  est  $\mathcal{B}$ -mesurable, alors  $\mathbb{E}[X \mid \mathcal{B}] = X$   $\mu$ -presque partout. L'application  $X \mapsto \mathbb{E}[X \mid \mathcal{B}]$  de  $\mathbb{L}^1(\Omega, \mathcal{A}, \mu)$  dans  $\mathbb{L}^1(\Omega, \mathcal{B}, \mu)$  est linéaire.



**Démonstration.** Posons  $g_n = \sup \{0, f, f + f \circ \phi, \dots, f + f \circ \phi + \dots + f \circ \phi^{n-1}\}$  et  $E_n = \{x \in X : g_n(x) > 0\}$ . Sur  $E_n$ , nous avons  $g_n = f + g_{n-1} \circ \phi$ . Sur  ${}^c E_n$ , nous avons  $g_n = 0$ . Sur  $X$ , nous avons  $g_{n-1} \circ \phi \geq 0$ . Donc

$$\begin{aligned} \int_{E_n} f \, d\mu &= \int_{E_n} g_n \, d\mu - \int_{E_n} g_{n-1} \circ \phi \, d\mu \\ &\geq \int_X g_n \, d\mu - \int_X g_{n-1} \circ \phi \, d\mu = \int_X (g_n - g_{n-1}) \, d\mu \geq 0. \end{aligned}$$

Or  $\{x \in X : f^*(x) > 0\} = \bigcup_{n \geq 1} E_n$ . Donc  $\int_{\{x \in X : f^*(x) > 0\}} f \, d\mu \geq 0$ .  $\square$

(1) Pour montrer l'existence de la limite, puisque  $\mathbb{Q}$  est dénombrable et dense dans  $\mathbb{R}$ , il suffit de voir que pour tous les rationnels  $\alpha, \beta$  tels que  $\alpha < \beta$ , l'ensemble

$$E_{\alpha, \beta} = \{x \in X : \liminf_{n \rightarrow \infty} S_n f(x) < \alpha < \beta < \limsup_{n \rightarrow \infty} S_n f(x)\}$$

est de mesure nulle. Remarquons que  $\mu(\phi^{-1}(E_{\alpha, \beta}) \Delta E_{\alpha, \beta}) = 0$  (voir le même argument que pour l'assertion (2) ci-dessous). Comme  $\limsup_{n \rightarrow \infty} S_n f \leq f^*$ , nous avons

$$E_{\alpha, \beta} \cap \{(f - \beta)^* = f^* - \beta > 0\} = E_{\alpha, \beta}.$$

De même, comme  $\liminf_{n \rightarrow \infty} S_n f \geq -(-f)^*$ , nous avons  $E_{\alpha, \beta} \cap \{(\alpha - f)^* > 0\} = E_{\alpha, \beta}$ .

Le lemme ergodique maximal appliqué sur la partie  $E_{\alpha, \beta}$  aux fonctions  $f - \beta$  et  $\alpha - f$  donne

$$\int_{E_{\alpha, \beta}} (f - \beta) \, d\mu \geq 0 \quad \text{et} \quad \int_{E_{\alpha, \beta}} (\alpha - f) \, d\mu \geq 0.$$

D'où  $\int_{E_{\alpha, \beta}} (\alpha - \beta) \, d\mu \geq 0$ , et donc  $\mu(E_{\alpha, \beta}) = 0$ .

(2) C'est clair, car la suite  $\frac{n+1}{n} S_{n+1} f(x) - S_n f(\phi(x)) = \frac{1}{n} f(x)$  converge presque partout vers 0, et  $f$  est finie presque partout car intégrable.

(3) Nous pouvons supposer que  $f$  est positive. Le résultat découle alors du lemme de Fatou (et du fait que  $f$  préserve  $\mu$ ) :

$$\int_X \tilde{f} \, d\mu = \int_X \limsup_{n \rightarrow \infty} S_n f \, d\mu \leq \limsup_{n \rightarrow \infty} \int_X S_n f \, d\mu = \int_X f \, d\mu.$$

(4) Lorsque  $f$  est bornée, cela découle du théorème de convergence dominée de Lebesgue, car si  $M > 0$  est tel que  $|f| \leq M$ , alors  $|S_n f| \leq M$  pour tout  $n \in \mathbb{N}$ . Le cas général s'en déduit par densité. En effet, pour tout  $\epsilon > 0$ , nous pouvons trouver  $g$  bornée telle que  $\|f - g\|_{\mathbb{L}^1} \leq \frac{\epsilon}{3}$ . Nous avons donc, pour tout  $n \geq 1$ , l'inégalité  $\|S_n f - S_n g\|_{\mathbb{L}^1} \leq \frac{\epsilon}{3}$  et par (3), nous avons  $\|\tilde{f} - \tilde{g}\|_{\mathbb{L}^1} \leq \frac{\epsilon}{3}$ . Pour  $n$  grand, nous avons alors  $\|S_n g - \tilde{g}\|_{\mathbb{L}^1} \leq \frac{\epsilon}{3}$  et donc  $\|S_n f - \tilde{f}\|_{\mathbb{L}^1} \leq \epsilon$ . C'est ce que nous voulions.

(5) Cela résulte de (4) appliqué à  $f|_A$ , car  $\int_A f = \int_A S_n f$ .

(6) Par (2), l'application  $\tilde{f}$  est constante presque partout, et  $\mu$  est une probabilité.  $\square$

Voici deux applications élémentaires du théorème ergodique de Birkhoff, voir aussi l'exercice E.16. Pour tout réel  $x$  dans  $[0, 1[$ , notons  $0, a_1 \dots a_n \dots$  son écriture décimale (si nous avons le choix, nous prenons l'écriture avec  $a_n$  nul pour  $n$  assez grand).

**Corollaire 2.4.** *Pour presque tout  $x$  dans  $[0, 1[$  pour la mesure de Lebesgue, pour tout chiffre  $j$  dans  $\{0, \dots, 9\}$ , « la proportion de  $j$  dans l'écriture décimale de  $x$  est égale à  $\frac{1}{10}$  » c'est-à-dire*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{Card}\{k \leq n, a_k = j\} = \frac{1}{10}.$$

**Démonstration.** Prenons  $X = [0, 1[$ ,  $\mu = dx$ ,  $\phi : x \rightarrow 10x \bmod 1$  (qui préserve la mesure de Lebesgue, en appliquant la proposition 0.2 (2) au  $\pi$ -système des intervalles semi-ouverts de  $[0, 1[$  d'extrémités 10-adiques), et  $f = \mathbb{1}_{[\frac{j}{10}, \frac{j+1}{10}]}$ . Nous avons alors

$$S_n f(x) = \frac{1}{n} \text{Card}\{k \leq n, a_k = j\} \xrightarrow{n \rightarrow \infty} \int_X f d\mu = \frac{1}{10},$$

car  $\phi$  est ergodique.<sup>56</sup> □

**Théorème 2.5. (Loi forte des grands nombres de Kolmogorov<sup>57</sup>)** *Soit  $(X_i)_{i \in \mathbb{N}}$  une suite de variables aléatoires réelles indépendantes, uniformément distribuées, et intégrables. Alors  $\frac{1}{n} \sum_{k=0}^{n-1} X_k$  converge presque sûrement vers  $\mathbb{E}[X_0]$ .*

**Démonstration.** Soit  $(\tilde{\Omega}, \tilde{\mathcal{B}}, \mathbb{P})$  l'espace de probabilité commun de définition des  $X_k$  pour  $k \in \mathbb{N}$  et  $\mu = (X_0)_* \mathbb{P}$  leur loi commune sur  $\mathbb{R}$ . Soient  $\sigma_+ : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$  le décalage à gauche sur  $\mathbb{R}$ , et  $\Theta : \tilde{\Omega} \rightarrow \mathbb{R}^{\mathbb{N}}$  l'application  $\omega \mapsto (X_i(\omega))_{i \in \mathbb{N}}$ . Notons  $\text{pr}_k : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$  la  $k$ -ème projection. Alors  $X_k = \text{pr}_k \circ \Theta$  et  $\text{pr}_k = \text{pr}_0 \circ \sigma_+^k$ . Notons

$$K = \{\omega \in \tilde{\Omega} : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} X_k(\omega) = \mathbb{E}[X_0]\}$$

et

$$K' = \{x \in \mathbb{R}^{\mathbb{N}} : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \text{pr}_0 \circ \sigma_+^k(x) = \mathbb{E}[X_0]\}.$$

Alors  $K = \Theta^{-1}(K')$ . Puisque le décalage à gauche  $\sigma_+$ , qui laisse invariante la mesure produit  $\mu^{\mathbb{N}}$  par l'exemple (2) de la partie 0.2, est ergodique<sup>58</sup> pour  $\mu^{\mathbb{N}}$ , et puisque

$$\int_{\mathbb{R}^{\mathbb{N}}} \text{pr}_0 d\mu^{\mathbb{N}} = \int_{\mathbb{R}} t d\mu(t) = \mathbb{E}[X_0],$$

---

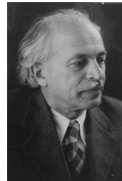
56. Nous reviendrons sur ce point dans la partie 4.5 (voir aussi l'exercice E.47), car  $\phi$  admet des propriétés dynamiques bien plus fortes que l'ergodicité. Soit  $f \in \mathcal{L}^2([0, 1])$  telle que  $f \circ \phi = f$  presque partout. Pour tout  $k \in \mathbb{Z}$ , soit  $c_k(f) = \int_0^1 e^{-2i\pi kt} f(t) dt$  le  $k$ -ème coefficient de Fourier de  $f$ . Alors par changement de variable, nous avons  $c_k(f) = c_{10^n k}(f \circ \phi^n) = c_{10^n k}(f)$ , qui tend vers 0 si  $k \neq 0$  quand  $n$  tend vers  $+\infty$  par le théorème de Parseval. Donc  $f = c_0(f)$  presque partout par la transformation de Fourier inverse, et  $f$  est constante presque partout.



Andrei Kolmogorov  
1903-1987



John Kingman  
1939-



Michael Fekete  
1886-1957



Hillel Furstenberg  
1935-



Harry Kesten  
1931-2019

58. Il possède des propriétés dynamiques bien plus fortes que l'ergodicité, nous y reviendrons par exemple dans la partie 3, voir la proposition 3.2.

nous avons  $\mu^{\mathbb{N}}(K') = 1$  par le théorème ergodique de Birkhoff 2.2 (6). Puisque les variables aléatoires  $X_i$  sont indépendantes, nous avons  $\Theta_*\mathbb{P} = \mu^{\mathbb{N}}$ . Donc  $\mathbb{P}(K) = \Theta_*\mathbb{P}(K') = \mu^{\mathbb{N}}(K') = 1$ . Ceci montre le résultat.  $\square$

### 2.3 Le théorème ergodique sous-additif de Kingman <sup>57</sup>

Le théorème suivant est un renforcement du théorème de Birkhoff, qui est tout autant utile, et servira en particulier dans la partie 2.4.

Il en existe de très nombreuses démonstrations, outre celle originelle de Kingman, données par Burkholder (1973), Derriennic (1975), Smeltzer (1977), Katznelson-Weiss (1982) et Neveu (1983). Nous donnerons celle d'approche algorithmique de Steele [Ste] ainsi que celle démontrant en même temps le théorème ergodique de Birkhoff due à Avila-Bochi (non publiée, voir [AvB]). Nous renvoyons à ces deux derniers articles pour d'autres références historiques.

**Théorème 2.6. (Théorème ergodique sous-additif de Kingman)** *Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret. Soit  $(f_n)_{n \in \mathbb{N} \setminus \{0\}}$  une suite de fonctions réelles mesurables, telle que  $f_1^+ = \max\{0, f_1\}$  appartienne à  $\mathbb{L}^1(X, \mu)$  et qui est  $\phi$ -sous-additive, c'est-à-dire vérifie*

$$\forall m, n \in \mathbb{N} \setminus \{0\}, \quad f_{m+n} \leq f_m + f_n \circ \phi^m .$$

Soit

$$\gamma = \lim_{n \rightarrow +\infty} \frac{1}{n} \int_X f_n d\mu = \inf_{n \in \mathbb{N} \setminus \{0\}} \frac{1}{n} \int_X f_n d\mu \in [-\infty, +\infty[ .$$

Alors il existe une fonction  $\widehat{f} : X \rightarrow [-\infty[$  mesurable et  $\mu$ -presque partout  $\phi$ -invariante, telle que  $\mu$ -presque partout, nous ayons la convergence

$$\lim_{n \rightarrow +\infty} \frac{1}{n} f_n = \widehat{f} . \tag{10}$$

De plus, la fonction  $(\widehat{f})^+$  est  $\mu$ -intégrable, et

$$\int_X \widehat{f} d\mu = \gamma .$$

Si  $\mu$  est ergodique, alors  $\widehat{f}$  est  $\mu$ -presque partout constante, égale à  $\gamma$ . Si  $\gamma > -\infty$ , alors pour tout  $n \in \mathbb{N}$ , les fonctions  $f_n$  et  $\widehat{f}$  sont intégrables, nous avons  $\|\widehat{f}\|_{\mathbb{L}^1} \leq \|f_1\|_{\mathbb{L}^1}$  et la convergence ci-dessus est aussi vérifiée dans  $\mathbb{L}^1(X, \mu)$ .

Les assertions (1), (2), (3) et (4) du théorème de Birkhoff 2.2, sous l'hypothèse supplémentaire que  $\mu$  est une mesure de probabilité, découlent de ce théorème. En effet, soit  $f \in \mathbb{L}^1(X, \mu)$ . Quitte à considérer les parties réelles  $\operatorname{Re} f$  et imaginaires  $\operatorname{Im} f$  de  $f$ , nous pouvons considérer que  $f$  est réelle. Appliquons le théorème de Kingman 2.6 à la suite pour  $n \in \mathbb{N} \setminus \{0\}$  des moyennes de Birkhoff non renormalisées <sup>59</sup>  $f_n = \sum_{k=0}^{n-1} f \circ \phi^k$  de  $f$ . Ceci est possible car cette suite vérifie la formule d'additivité exacte  $f_{m+n} = f_m + f_n \circ \phi^m$  pour tous les  $m, n \in \mathbb{N} \setminus \{0\}$ , ainsi que l'intégrabilité de  $f_1^+ = f^+$ . Pour tout  $n \in \mathbb{N}$ , l'invariance de  $\mu$  montre que  $\frac{1}{n} \int_X f_n d\mu = \int_X f d\mu$ , donc la constante  $\gamma$  de la suite  $(f_n)_{n \in \mathbb{N}}$  vaut

59. parfois appelées les sommes de Birkhoff

$\int_X f d\mu$ . Les assertions (1) à (4) du théorème de Birkhoff 2.2 découlent alors du théorème de Kingman 2.6 car  $S_n f = \frac{1}{n} f_n$  (de sorte que  $\tilde{f} = \hat{f}$  avec les notations de ces deux énoncés...). Nous utiliserons toutefois le théorème ergodique de Birkhoff 2.2 dans la première démonstration ci-dessous, voir la fin de cette partie pour une seconde démonstration qui ne l'utilise pas.

L'hypothèse que seulement  $f_1^+$  est intégrable est par exemple impliquée par l'hypothèse que  $f_1$  elle-même est intégrable, et elle sera utilisée par exemple pour le théorème d'Oseledets 2.12.

**Démonstration.** Par la sous-additivité et par récurrence, pour tout  $n \geq 1$ , nous avons  $f_n \leq f_1 + f_1 \circ \phi + \dots + f_1 \circ \phi^{n-1}$ , donc  $0 \leq f_n^+ \leq f_1^+ + f_1^+ \circ \phi + \dots + f_1^+ \circ \phi^{n-1}$ . Puisque  $f_1^+$  est intégrable et par l'invariance de  $\mu$ , l'application  $f_n^+$  est intégrable. Rappelons que si  $g : X \rightarrow \mathbb{R}$  est une fonction mesurable et si  $g^+$  est intégrable pour  $\mu$ , alors<sup>60</sup>

$$\int_X g d\mu = \int_X g^+ d\mu - \int_X (-g)^+ d\mu$$

est bien définie dans  $[-\infty, +\infty[$ . En particulier, l'intégrale  $\int_X f_n d\mu$  est bien définie dans  $[-\infty, +\infty[$ , et si  $\gamma > +\infty$ , alors comme  $\int_X f_n d\mu \geq n\gamma$ , la fonction  $f_n$  est intégrable pour tout  $n \geq 1$ .

Par l'invariance de  $\mu$  par  $\phi$ , nous avons

$$\int_X f_{m+n} d\mu \leq \int_X f_m d\mu + \int_X f_n d\mu$$

pour tous les  $m, n \in \mathbb{N} \setminus \{0\}$ . Le fait que la limite dans la première définition de  $\gamma$  dans l'énoncé ci-dessus existe, et que les deux définitions de  $\gamma$  coïncident, est donc une conséquence du lemme classique de sous-additivité suivant appliqué à la suite  $(a_n = \int_X f_n d\mu)_{n \in \mathbb{N}}$  dans  $[-\infty, +\infty[$ . Il est connu sous le nom de lemme de Fekete<sup>57</sup>, et il sera aussi utile pour la définition de l'entropie mesurée des transformations (voir la partie 5.2). Ce lemme est en fait une conséquence du théorème ergodique de Kingman appliqué à l'unique système dynamique probabilisé dont l'espace des phases est réduit à un point (qui est ergodique)!

**Lemme 2.7. (Lemme de Fekete)** Soit  $(a_n)_{n \geq 1}$  une suite dans  $[-\infty, +\infty[$  telle que, pour tous les  $m, n \geq 1$ , nous ayons  $a_{m+n} \leq a_m + a_n$ . Alors la limite  $\lim_{n \rightarrow \infty} \frac{a_n}{n}$  existe dans  $[-\infty, +\infty[$  et vaut  $\inf_{n \geq 1} \frac{a_n}{n}$ .

**Démonstration.** S'il existe  $k \geq 1$  tel que  $a_k = -\infty$ , alors par sous-additivité, nous avons  $a_n = -\infty$  pour tout  $n \geq 1$ , et le résultat est immédiat. Supposons donc que  $a_n \in \mathbb{R}$  pour tout  $n \geq k$ . Soit  $\ell = \inf_{n \geq 1} \frac{a_n}{n}$ .

Supposons tout d'abord  $\ell > -\infty$ . Alors la suite positive  $(b_n = a_n - \ell n)_{n \in \mathbb{N}}$  est encore sous-additive, et pour tout  $\epsilon > 0$ , il existe  $n_0 \geq 1$  tel que  $\frac{b_{n_0}}{n_0} \leq \epsilon$ . Écrivons par division euclidienne  $n = qn_0 + r$  avec  $r < n_0$ . Par sous-additivité, pour tout  $n \geq n_0$ , nous avons alors  $\frac{b_n}{n} \leq \frac{qb_{n_0} + rb_1}{n} \leq \frac{b_{n_0}}{n_0} + \frac{n_0 b_1}{n}$ . Donc  $0 \leq \frac{b_n}{n} \leq 2\epsilon$  pour  $n$  grand et  $\lim_{n \geq 1} \frac{b_n}{n} = 0$ .

Si maintenant  $\ell = -\infty$ , montrons de manière similaire que la suite  $(\frac{a_n}{n})_{n \in \mathbb{N}}$  converge vers  $-\infty$ . Pour tout  $A > 0$ , il existe  $n_0 \geq 1$  tel que  $\frac{a_{n_0}}{n_0} \leq -2A - 1$ . Donc pour tout  $n \geq n_0$

60. La convention standard est que  $-\infty + a = -\infty$  pour tout  $a \in [-\infty, +\infty[$ .

assez grand, nous avons  $\frac{a_n}{n} \leq \frac{qa_{n_0} + ra_1}{n} \leq \frac{a_{n_0}}{n_0} \left(1 - \frac{n_0}{n}\right) + \frac{n_0|a_1|}{n} \leq -A$ , ce qui montre le résultat.  $\square$

Dans cette première démonstration du théorème de Kingman, nous ne démontrerons le résultat que lorsque la suite  $(f_n)_{n \geq 1}$  est dans  $\mathbb{L}^1(X, \mu)$ , ce qui est le cas dans la plupart des applications, en renvoyant par exemple à [Via, Theo. 3.3] pour le passage au cas général, ainsi qu'à la seconde démonstration ci-dessous.

Montrons tout d'abord le théorème de Kingman 2.6 lorsque  $(f_n)_{n \geq 1}$  est de plus supposée être une suite de fonctions négatives. Le résultat général s'y ramène en appliquant ce cas particulier à la suite  $(g_n)_{n \geq 1}$  où  $g_n = f_n - \sum_{i=1}^{n-1} f_1 \circ \phi^i$ . En effet, cette suite est encore dans  $\mathbb{L}^1(X, \mu)$ , elle est négative par la sous-additivité de  $(f_n)_{n \geq 1}$ , et elle reste sous-additive car les moyennes de Birkhoff non normalisées sont additives. Supposons le théorème de Kingman vérifié pour  $(g_n)_{n \geq 1}$ . Notons  $\hat{g} = \lim_{n \rightarrow +\infty} \frac{1}{n} g_n$  presque partout, et  $\gamma' = \lim_{n \rightarrow +\infty} \frac{1}{n} \int_X g_n d\mu = \int_X \hat{g} d\mu$ . Comme  $\gamma = \gamma' + \int_X f_1 d\mu$ , nous avons  $\gamma > -\infty$  si et seulement si  $\gamma' > -\infty$ . Alors  $\frac{1}{n} f_n = \frac{1}{n} g_n + S_n(f_1)$  converge presque partout, et dans  $\mathbb{L}^1(X, \mathcal{B}, \mu)$  si  $\gamma > -\infty$ , par le théorème de Birkhoff, vers  $\hat{f} = \hat{g} + \tilde{f}_1$ , où  $\tilde{f}_1 = \lim_{n \rightarrow +\infty} S_n(f_1)$  (égale à l'espérance conditionnelle  $\mathbb{E}(f_1 | \mathcal{I})$  en rappelant que  $\mathcal{I}$  est la tribu des invariants de  $\phi$ ). Enfin,

$$\gamma = \lim_{n \rightarrow +\infty} \frac{1}{n} \int_X f_n d\mu = \lim_{n \rightarrow +\infty} \frac{1}{n} \int_X g_n d\mu + \int_X f_1 d\mu = \int_X (\hat{g} + \tilde{f}_1) d\mu = \int_X \hat{f} d\mu,$$

toujours par le théorème de Birkhoff pour l'avant-dernière égalité.

Considérons les applications mesurables  $\underline{f} = \liminf_{n \rightarrow +\infty} \frac{1}{n} f_n$  et  $\bar{f} = \limsup_{n \rightarrow +\infty} \frac{1}{n} f_n$ . Montrons que  $\underline{f}$  est presque partout  $\phi$ -invariante. En effet, nous avons  $\underline{f} \leq \underline{f} \circ \phi$  par division par  $n$  et passage à la limite inférieure de l'inégalité de sous-additivité  $f_{n+1} \leq f_1 + f_n \circ \phi$ . Donc pour tout  $\alpha \in \mathbb{R}$ , nous avons  $A_\alpha = \{x \in X : \underline{f} \leq \alpha\} \subset \{x \in X : \underline{f} \circ \phi \leq \alpha\} = \phi^{-1}(A_\alpha)$ . Si  $\underline{f}$  et  $\underline{f} \circ \phi$  n'étaient pas presque partout égales, il existerait  $\alpha \in \mathbb{Q}$  tel que  $\mu(A_\alpha) < \mu(\phi^{-1}(A_\alpha))$ , ce qui contredirait l'invariance par  $\phi$  de la mesure  $\mu$ . Nous pouvons donc supposer que  $\underline{f} = \underline{f} \circ \phi$ . De même nous pouvons supposer que  $\bar{f} = \bar{f} \circ \phi$ .

Montrons l'existence de la limite dans la formule (10). Soient  $\epsilon > 0$ ,  $N \in \mathbb{N} \setminus \{0\}$  et  $M \in ]0, +\infty[$ . Posons  $F_M = \max\{-M, \underline{f}\}$  et considérons l'ensemble mesurable

$$B_{N,M} = \{x \in X : \forall \ell \in \llbracket 1, N \rrbracket, \frac{1}{\ell} f_\ell(x) > F_M(x) + \epsilon\}.$$

Pour tous les point  $x \in X$  et les entiers  $n \geq N$ , nous allons décomposer algorithmiquement l'intervalle entier  $\llbracket 1, n-1 \rrbracket$  comme la réunion disjointe de certains sous-intervalles entiers, en parcourant dans l'ordre ses éléments.

Soit  $k$  le plus petit élément de  $\llbracket 1, n-1 \rrbracket$  qui n'est pas dans un des sous-intervalles déjà construits (et en particulier  $k = 1$  à la première étape). Si  $\phi^k(x) \notin B_{N,M}$ , considérons  $\ell \in \llbracket 1, N \rrbracket$  le plus petit entier tel que  $\frac{1}{\ell} f_\ell(\phi^k(x)) \leq F_M(\phi^k(x)) + \epsilon = F_M(x) + \epsilon$ . Si  $k + \ell \leq n$ , nous prenons  $\llbracket k, k + \ell - 1 \rrbracket$  comme l'un de nos sous-intervalles. Si  $k + \ell > n$  ou si  $\phi^k(x) \in B_{N,M}$ , nous prenons le singleton  $\{k\}$  comme l'un de nos sous-intervalles.

Ainsi pour tout  $x \in X$ , nous avons écrit  $\llbracket 1, n-1 \rrbracket$  comme la réunion disjointe de sous-intervalles de trois type,

- une famille  $(\llbracket \tau_i, \tau_i + \ell_i - 1 \rrbracket)_{1 \leq i \leq u}$  où  $f_{\ell_i}(\phi^{\tau_i}(x)) \leq (F_M(x) + \epsilon)\ell_i$ ,
- une famille de singletons  $(\{\sigma_i\})_{1 \leq i \leq v}$  tels que  $\mathbb{1}_{B_{N,M}}(\phi^{\sigma_i}(x)) = 1$ ,
- une famille de singletons  $(\{\omega_i\})_{1 \leq i \leq w}$  contenus dans  $\llbracket n - N + 1, n - 1 \rrbracket$ .

Par la propriété de sous-additivité appliquée à cette décomposition, et puisque la fonction  $f_1$  est négative ou nulle, nous avons

$$\begin{aligned} f_n(x) &\leq \sum_{i=1}^u f_{\ell_i}(\phi^{\tau_i}(x)) + \sum_{i=1}^v f_1(\phi^{\sigma_i}(x)) + \sum_{i=1}^w f_1(\phi^{\omega_i}(x)) \\ &\leq (F_M(x) + \epsilon) \sum_{i=1}^u \ell_i \leq F_M(x) \sum_{i=1}^u \ell_i + n\epsilon. \end{aligned} \quad (11)$$

Puisque les sous-intervalles ci-dessus recouvrent  $\llbracket 1, n-1 \rrbracket$ , nous avons

$$n - \sum_{k=1}^n \mathbb{1}_{B_{N,M}}(\phi^k(x)) - N \leq \sum_{i=1}^u \ell_i.$$

Donc par le théorème ergodique de Birkhoff et avec les notations de son énoncé, nous avons  $\mu$ -presque partout

$$\liminf_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^u \ell_i \geq \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n (1 - \mathbb{1}_{B_{N,M}})(\phi^k(x)) = 1 - \widetilde{\mathbb{1}_{B_{N,M}}}.$$

Par la formule (11), nous avons donc  $\mu$ -presque partout

$$\bar{f}(x) = \limsup_{n \rightarrow +\infty} \frac{1}{n} f_n(x) \leq F_M(x)(1 - \widetilde{\mathbb{1}_{B_{N,M}}}) + \epsilon.$$

En fixant  $M$  et en faisant tendre  $N$  vers  $+\infty$ , puisque la suite d'applications  $(\mathbb{1}_{B_{N,M}})_{N \in \mathbb{N}}$ , qui sont mesurables et uniformément bornées, converge  $\mu$ -presque partout vers 0 par la définition de  $B_{N,M}$  et par la continuité de l'application  $h \mapsto \tilde{h} = \mathbb{E}(h \mid \mathcal{I})$  dans  $\mathbb{L}^1(X, \mu)$  vue par exemple dans le théorème de Birkhoff 2.2 (3), nous obtenons que  $\mu$ -presque partout

$$\bar{f}(x) \leq F_M(x) + \epsilon.$$

En faisant tendre  $M$  vers  $+\infty$ , puis  $\epsilon$  vers 0 (dans  $\mathbb{Q}$ ), nous obtenons donc  $\mu$ -presque partout

$$\bar{f}(x) \leq \underline{f}(x),$$

ce qui montre la convergence cherchée (par le critère d'égalité des  $\limsup$  et  $\liminf$ ).

Par la sous-additivité, pour tout  $n \in \mathbb{N}$ , nous avons  $f_n \leq \sum_{i=0}^{n-1} f_1 \circ \phi^i \leq \sum_{i=0}^{n-1} f_1^+ \circ \phi^i$ . Donc  $\frac{1}{n} f_n^+ \leq S_n f_1^+$ . L'intégrabilité de  $(\hat{f})^+$  découle du théorème de Birkhoff, qui dit en particulier que puisque l'application  $f_1^+$  est intégrable, la limite presque partout de ses moyennes de Birkhoff  $S_n f_1^+$  l'est aussi.

L'avant-dernière affirmation du théorème 2.6 découle de l'invariance presque partout de  $\hat{f}$  et de l'équivalence des définitions 2.1 de l'ergodicité.

Si  $\gamma > +\infty$ , alors  $\int_X |\hat{f}| d\mu = \int_X \hat{f}^+ d\mu + \int_X (-\hat{f})^+ d\mu = 2\|\hat{f}^+\|_{\mathbb{L}^1} - \gamma < +\infty$ . La dernière affirmation du théorème 2.6 découle du théorème de convergence dominée de Lebesgue.  $\square$

**Autre démonstration du théorème ergodique sous-additif de Kingman 2.6.** Nous donnons maintenant une autre démonstration du théorème 2.6, due à Avila et Bochi [AvB],

qui au contraire de la précédente, d'une part n'utilise que l'hypothèse optimale d'intégrabilité de  $f_1^+$ , et d'autre part, n'utilise pas le théorème ergodique de Birkhoff, qui en est donc une conséquence, du moins lorsque la mesure  $\mu$  est une mesure de probabilité, comme expliqué dans la remarque suivant l'énoncé du théorème 2.6.

Soient  $(X, \mathcal{B}, \mu, \phi)$  et  $(f_n)_{n \geq 1}$  comme dans l'énoncé du théorème 2.6. Comme dans le début de la démonstration précédente, les fonctions  $f_n^+$  sont intégrables, donc les intégrales  $\int_X f_n d\mu$  sont bien définies dans  $[-\infty, +\infty[$ . Par la sous-additivité de  $(f_n)_{n \in \mathbb{N}}$ , l'invariance de  $\mu$  et le lemme de Fekete, la limite  $\lim_{n \rightarrow +\infty} \frac{1}{n} \int_X f_n d\mu$  existe et vaut  $\gamma = \inf_{n \geq 1} \frac{1}{n} \int_X f_n d\mu \in [-\infty, +\infty[$ . Considérons de nouveau les applications mesurables  $\underline{f} = \liminf_{n \rightarrow +\infty} \frac{1}{n} f_n$  et  $\bar{f} = \limsup_{n \rightarrow +\infty} \frac{1}{n} f_n$ . Comme dans la démonstration précédente, par la sous-additivité, nous avons  $\underline{f} = \underline{f} \circ \phi$  presque partout et  $\bar{f} = \bar{f} \circ \phi$  presque partout.

**Lemme 2.8.** *Nous avons  $\int_X \underline{f} d\mu = \gamma$ .*

**Démonstration. Étape 1.** Supposons tout d'abord qu'il existe  $C > 0$  tel que  $\frac{f_n}{n} \geq -C$  pour tout  $n \geq 1$ .

En particulier, l'application  $f_1$ , minorée par  $-C$  sur un espace de probabilité et de partie positive intégrable, est intégrable. Par le lemme de Fatou appliqué aux fonctions positives  $\frac{f_n}{n} + C$ , nous avons

$$\int_X \underline{f} d\mu = \int_X \liminf_{n \rightarrow +\infty} \frac{f_n}{n} d\mu \leq \liminf_{n \rightarrow +\infty} \frac{1}{n} \int_X f_n d\mu = \gamma < +\infty, \quad (12)$$

et en particulier, comme  $\underline{f}$  est minorée par  $-C$  sur un espace de probabilité, l'application  $\underline{f}$  est intégrable. Fixons  $\epsilon > 0$ . Considérons la famille croissante  $(E_k)_{k \geq 1}$  de parties mesurables de  $X$  où

$$E_k = \{x \in X : \exists j \in \llbracket 1, k \rrbracket, \frac{f_j(x)}{j} \leq \underline{f}(x) + \epsilon\}.$$

Notons que  $X = \bigcup_{k \geq 1} E_k$  par la définition de  $\underline{f}$ . Considérons la fonction  $g_k : X \rightarrow \mathbb{R}$  définie par  $g_k = \begin{cases} \underline{f} + \epsilon & \text{sur } E_k \\ f_1 & \text{sur } {}^c E_k, \end{cases}$  qui est intégrable.

Montrons que pour presque tout  $x \in X$ , et tous les entiers  $n, k$  tels que  $n \geq k \geq 1$  nous avons

$$f_n(x) \leq \sum_{i=0}^{n-k-1} g_k(\phi^i(x)) + \sum_{i=n-k}^{n-1} \max\{g_k, f_1\}(\phi^i(x)). \quad (13)$$

En effet, pour presque tout  $x \in X$ , la fonction  $\underline{f}$  est constante sur l'orbite positive  $\mathcal{O}^+(x)$  de  $x$  par invariance presque partout. Pour tout tel  $x$  et  $k \geq 1$ , construisons par récurrence des entiers  $m_0 \leq n_1 < m_1 \leq n_2 < m_2 \leq \dots$  (en nombre fini ou infini) de la manière suivante. Posons  $m_0 = 0$ . Soit  $j \geq 1$  et supposons  $m_{j-1}$  construit. Si l'ensemble  $\{\ell \geq m_{j-1} : \phi^\ell(x) \in E_k\}$  est vide, la construction s'arrête au rang  $j - 1$ . Sinon, posons  $n_j = \min\{\ell \geq m_{j-1} : \phi^\ell(x) \in E_k\}$ . Par la définition de  $E_k$ , soit  $m_j$  un entier tel que  $m_j - n_j \in \llbracket 1, k \rrbracket$  et

$$f_{m_j - n_j}(\phi^{n_j}(x)) \leq (m_j - n_j)(\underline{f}(\phi^{n_j}(x)) + \epsilon). \quad (14)$$

Soit  $n \geq k$ , et soit  $\ell$  le plus grand entier tel que  $m_\ell$  soit défini et  $m_\ell \leq n$ . Par la sous-additivité et par récurrence, nous avons

$$f_n(x) \leq \sum_{j=1}^{\ell} f_{m_j - n_j}(\phi^{n_j}(x)) + \sum_{i \in \bigcup_{j=0}^{\ell-1} \llbracket m_j, n_{j+1} \rrbracket \cup \llbracket m_\ell, n \rrbracket} f_1(\phi^i(x)). \quad (15)$$

Par convention, posons  $\min\{n_{\ell+1}, n\} = n$  si l'élément  $n_{\ell+1}$  n'est pas défini (il convient de poser alors  $n_{\ell+1} = +\infty$ ). Pour tout entier  $i \in \bigcup_{j=0}^{\ell-1} \llbracket m_j, n_{j+1} \rrbracket \cup \llbracket m_\ell, \min\{n_{\ell+1}, n\} \rrbracket$ , nous avons alors  $f_1(\phi^i(x)) = g_k(\phi^i(x))$  car  $\phi^i(x) \notin E_k$  par construction. De plus, par la formule (14), par l'invariance de  $\underline{f}$  le long de l'orbite positive de  $x$ , et comme nous avons  $g_k \geq \underline{f} + \epsilon$  sur tout  $X$  (car pour tout  $x \in {}^c E_k$ , nous avons  $\frac{f_1(x)}{1} > \underline{f}(x) + \epsilon$  par la définition de  $E_k$ ), nous obtenons

$$f_{m_j - n_i}(\phi^{n_j}(x)) \leq \sum_{i \in \llbracket n_j, m_j \rrbracket} (\underline{f}(\phi^i(x)) + \epsilon) \leq \sum_{i \in \llbracket n_j, m_j \rrbracket} g_k(\phi^i(x)).$$

Donc la formule (15) devient

$$f_n(x) \leq \sum_{i=1}^{\min\{n_{\ell+1}, n\}-1} g_k(\phi^i(x)) + \sum_{i=\min\{n_{\ell+1}, n\}}^{n-1} f_1(\phi^i(x)).$$

Ceci démontre la formule (13), car  $n_{\ell+1} > n - k$  puisque par la maximalité de  $\ell$ , nous avons  $n_{\ell+1} + k \geq m_{\ell+1} > n$  si  $n_{\ell+1}$  est défini et sinon  $n_{\ell+1} = +\infty$ .

En intégrant la formule (13) et par l'invariance de la mesure, nous avons donc

$$\int_X f_n \, d\mu \leq (n - k) \int_X g_k \, d\mu + k \int_X \max\{g_k, f_1\} \, d\mu. \quad (16)$$

En divisant par  $n$  et en faisant tendre  $n$  vers  $+\infty$ , nous avons donc  $\gamma \leq \int_X g_k \, d\mu$ . En faisant tendre  $k$  vers  $+\infty$ , comme  $g_k$  tend simplement vers  $\underline{f} + \epsilon$  puisque  $X = \bigcup_{k \in \mathbb{N}} E_k$ , et puisque  $g_k$  est minorée par  $-C$  et majorée par la fonction intégrable  $\max\{\underline{f} + \epsilon, f_1\}$ , par le théorème de convergence dominée de Lebesgue, nous avons  $\gamma \leq \int_X \underline{f} \, d\mu + \epsilon$ . En faisant tendre  $\epsilon$  vers 0, et comme  $\gamma \geq \int_X \underline{f} \, d\mu$  par la formule (12), ceci démontre le lemme 2.8 sous les hypothèses de l'étape 1.

**Étape 2.** Dédouons maintenant le cas général du lemme 2.8 à partir de l'étape 1, par un argument usuel de troncature inférieure.

Pour tout  $C > 0$ , posons  $f_n^C = \max\{f_n, -Cn\}$ . La suite  $(f_n^C)_{n \in \mathbb{N}}$  reste sous-additive, et  $(f_1^C)^+ = f_1^+$  est intégrable. Nous avons

$$\underline{f}^C = \liminf_{n \rightarrow +\infty} \frac{1}{n} f_n^C = \max\{\underline{f}, -C\} \quad \text{et} \quad \overline{f}^C = \limsup_{n \rightarrow +\infty} \frac{1}{n} f_n^C = \max\{\overline{f}, -C\}.$$

Par le théorème de convergence monotone en faisant tendre  $C$  vers  $+\infty$  pour la première et la dernière égalité, et par l'étape 1 pour la dernière égalité, nous obtenons donc

$$\int_X \underline{f} \, d\mu = \inf_{C > 0} \int_X \underline{f}^C \, d\mu = \inf_{C > 0} \inf_{n \geq 1} \int_X \frac{f_n^C}{n} \, d\mu = \inf_{n \geq 1} \inf_{C > 0} \int_X \frac{f_n^C}{n} \, d\mu = \inf_{n \geq 1} \int_X \frac{f_n}{n} \, d\mu = \gamma.$$

Ceci démontre le lemme 2.8. □



**Lemme 2.9.** Si  $g \in \mathbb{L}^1(X, \mu)$ , alors  $\frac{1}{n} g \circ \phi^n$  converge presque partout vers 0.

**Démonstration.** Soit  $N \geq 1$ . Si  $m \geq 1$ , posons  $A_m = \bigcup_{n \geq m} \{x \in X : |g(\phi^n(x))| \geq \frac{n}{N}\}$ . Montrons que  $\lim_{m \rightarrow +\infty} \mu(A_m) = 0$ . Ceci montrera que

$$\mu(\{x \in X : \limsup_{n \rightarrow +\infty} \frac{1}{n} |g(\phi^n(x))| \geq \frac{1}{N}\}) = 0,$$

et le résultat en découlera par intersection dénombrable. Or, par l'invariance de la mesure, et le fait que  $\frac{k+1}{k} \leq 2$  pour tout  $k \geq 1$ , nous avons

$$\begin{aligned} & \sum_{n \geq 1} \mu(\{x \in X : |g(\phi^n(x))| \geq \frac{n}{N}\}) = \sum_{n \geq 1} \mu(\{x \in X : |g(x)| \geq \frac{n}{N}\}) \\ &= \sum_{n \geq 1} \sum_{k \geq n} \mu(\{x \in X : k \leq N |g(x)| < k+1\}) \\ &= \sum_{k \geq 1} k \mu(\{x \in X : k \leq N |g(x)| < k+1\}) \leq 2 \int_X N |g(x)| d\mu < +\infty. \end{aligned}$$

Puisque la première série de la première ligne de cette formule converge, son reste tend vers 0, et le résultat cherché découle alors du fait que par sous-additivité nous avons  $\mu(A_m) \leq \sum_{n \geq m} \mu(\{x \in X : |g(\phi^n(x))| \geq \frac{n}{N}\})$ .  $\square$

**Lemme 2.10.** Pour tout  $k \geq 1$ , nous avons  $\limsup_{n \rightarrow +\infty} \frac{f_{kn}}{n} = k \limsup_{n \rightarrow +\infty} \frac{f_n}{n}$ .

**Démonstration.** Fixons  $k \geq 1$ . L'inégalité  $\leq$  est immédiate. Montrons celle opposée. Posons  $g = \max\{f_1^+, \dots, f_k^+\}$ . Nous avons vu au début de la démonstration précédente que  $f_\ell^+$  est intégrable pour tout  $\ell \geq 1$ . Donc  $g \in \mathbb{L}^1(X, \mu)$ .

Pour tout  $n \geq 1$ , écrivons  $n = mk + r$  la division euclidienne de  $n$  par  $k$ , avec  $m, r \in \mathbb{N}$  qui dépendent de  $n$  et  $r < k$ . Par la sous-additivité, nous avons

$$f_n \leq f_{mk} + f_r \circ \phi^{mk} \leq f_{mk} + f_r^+ \circ \phi^{mk} \leq f_{mk} + g \circ \phi^{mk}.$$

Puisque  $\lim_{n \rightarrow +\infty} \frac{mk}{n} = 1$  et  $g \in \mathbb{L}^1(X, \mu)$ , nous avons  $\lim_{n \rightarrow +\infty} \frac{1}{n} g \circ \phi^{mk} = 0$  par le lemme 2.9 précédent. Le résultat en découle en divisant par  $n$  l'inégalité  $f_n \leq f_{mk} + g \circ \phi^{mk}$  ci-dessus et en passant à la limite supérieure, puisque  $\lim_{n \rightarrow +\infty} \frac{m}{n} = \frac{1}{k}$ .  $\square$

**Fin de la démonstration du théorème de Kingman.** Considérons d'abord de nouveau le cas où il existe  $C > 0$  tel que  $\frac{f_n}{n} \geq -C$  pour tout  $n \geq 1$ . Montrons en utilisant les idées précédentes que

$$\int_X \bar{f} d\mu \leq \gamma. \tag{17}$$

Avec le lemme 2.8, ceci montrera que  $\int_X \bar{f} d\mu \leq \gamma \leq \int_X \underline{f} d\mu$ . Comme  $\underline{f} \leq \bar{f}$  et puisque ces fonctions sont intégrables, cela montrera que  $\underline{f} = \bar{f}$  presque partout. La convergence cherchée dans l'énoncé du théorème de Kingman en découlera (par le critère d'égalité des lim sup et lim inf). Les autres assertions du théorème de Kingman se démontrent comme dans la démonstration précédente.

Afin de montrer la formule (17), fixons  $k \geq 1$ . Pour tout  $n \geq 1$ , posons  $F_n = -\sum_{j=0}^{n-1} f_k \circ \phi^{jk}$ ,<sup>61</sup> et remarquons de nouveau que la suite  $(F_n)_{n \geq 1}$  est (exactement) additive pour  $\phi^k$ . De plus,  $F_1 = -f_k \leq Ck$ , donc  $F_1^+$  est intégrable. Par l'invariance de la mesure, nous avons  $\int_X \frac{F_n}{n} d\mu = -\int_X f_k d\mu$ . En posant  $\underline{F} = \liminf_{n \rightarrow +\infty} \frac{F_n}{n}$ , par le lemme 2.8, nous avons donc  $\int_X \underline{F} d\mu = \liminf_{n \rightarrow +\infty} \int_X \frac{F_n}{n} d\mu = -\int_X f_k d\mu$ . La sous-additivité de  $(f_n)_{n \geq 1}$  et le lemme 2.10 montrent que

$$-\underline{F} = \limsup_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_k \circ \phi^{jk} \geq \limsup_{n \rightarrow +\infty} \frac{f_{kn}}{n} = k \limsup_{n \rightarrow +\infty} \frac{f_n}{n} = k \bar{f}.$$

Donc en divisant par  $n$  et en intégrant, nous avons  $\int_X \bar{f} d\mu \leq -\frac{1}{k} \int_X \underline{F} d\mu = \frac{1}{k} \int_X f_k d\mu$ . En prenant la borne inférieure sur  $k \geq 1$ , ceci montre la formule (17) sous l'hypothèse ci-dessus de minoration des fonctions  $\frac{f_n}{n}$ .

Dans le cas général, la fin de la démonstration du théorème de Kingman procède de nouveau par l'argument usuel de troncature inférieure déjà employé. Pour tout  $C > 0$ , les fonctions  $\underline{f}^C$  et  $\bar{f}^C$ , définies dans l'étape 2 de la démonstration ci-dessus du lemme 2.8, coïncident presque partout par ce qui précède. Comme  $\lim_{C \rightarrow +\infty} \underline{f}^C = \underline{f}$  et  $\lim_{C \rightarrow +\infty} \bar{f}^C = \bar{f}$ , le résultat s'en déduit.  $\square$

## 2.4 Le théorème ergodique multiplicatif d'Oseledets

Notons  $\ln^+ = \max\{0, \ln\} : ]0, +\infty[ \rightarrow [0, +\infty[$ . Soient  $d \in \mathbb{N} \setminus \{0\}$ ,  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ <sup>62</sup> et  $\mathcal{M}_d(\mathbb{K})$  l'algèbre des matrices  $d \times d$  à coefficients dans  $\mathbb{K}$ , de matrice unité  $I_d$ . Fixons une norme quelconque  $\|\cdot\|$  sur l'espace vectoriel  $\mathcal{M}_d(\mathbb{K})$  sur  $\mathbb{K}$ .

Soit  $(X, \mathcal{B}, \phi)$  un système dynamique mesurable à temps discret. Fixons une application mesurable  $M : X \rightarrow \mathcal{M}_d(\mathbb{K})$ . Nous appellerons *produit dynamique de matrices* à gauche associée à  $(\phi, M)$  la suite  $(M_n)_{n \in \mathbb{N}}$  d'applications mesurables de  $X$  dans  $\mathcal{M}_d(\mathbb{K})$  définie par récurrence par  $M_0 = I_d$  et pour tout  $n \in \mathbb{N}$  par  $M_{n+1} = M \circ \phi^n \cdot M_n$  (en notant par  $\cdot$  le produit matriciel des applications à valeurs dans  $\mathcal{M}_d(\mathbb{K})$ , que nous omettrons dans la suite), de sorte que<sup>63</sup> (avec la convention usuelle qu'un produit vide d'éléments de  $\mathcal{M}_d(\mathbb{K})$  est  $I_d$ )

$$\forall n \in \mathbb{N}, \forall x \in X, \quad M_n(x) = M(\phi^{n-1}(x)) M(\phi^{n-2}(x)) \dots M(\phi(x)) M(x). \quad (18)$$

La suite  $(M_n)_{n \in \mathbb{N}}$  vérifie la propriété de *cocycle multiplicatif* à gauche pour  $\phi$

$$\forall m, n \in \mathbb{N}, \forall x \in X, \quad M_{m+n}(x) = M_n(\phi^m(x)) M_m(x). \quad (19)$$

Pour cette raison, l'application  $M$  est parfois appelée un *cocycle linéaire* pour le système dynamique  $(X, \mathcal{B}, \phi)$ .

61. qui est la  $n$ -ème somme de Birkhoff de la fonction  $-f_k$  pour la transformation  $\phi^k$

62. La plupart des résultats ci-dessous restent valable lorsque  $(\mathbb{K}, |\cdot|)$  est un corps commutatif valué localement compact, comme par exemple le corps  $\mathbb{Q}_p$  muni de sa valuation  $p$ -adique pour tout premier  $p$ .

63. Comme dans de nombreuses références, il est tout aussi intéressant (et lié par passage à l'inverse) d'étudier les comportements asymptotiques des produits dynamiques de matrices où les matrices sont multipliées dans l'autre ordre  $x \mapsto M(x)M(\phi(x)) \dots M(\phi^{n-1}(x))$ . Notre choix d'ordre vient du fait que nous allons faire agir les matrices (linéairement, à gauche) sur les vecteurs, voir la remarque (1) suivante.

Si le système dynamique  $(X, \mathcal{B}, \phi)$  est inversible et si  $M$  est à valeurs dans le groupe linéaire  $\text{GL}_d(\mathbb{K})$ , nous posons, pour tous les  $n \in \mathbb{N}$  et  $x \in X$ ,

$$\begin{aligned} M_{-n}(x) &= (M_n \circ \phi^{-n}(x))^{-1} \\ &= (M(\phi^{-n}(x)))^{-1} (M(\phi^{-n+1}(x)))^{-1} \dots (M(\phi^{-2}(x)))^{-1} (M(\phi^{-1}(x)))^{-1}. \end{aligned}$$

Il est alors immédiat de montrer par récurrence que la formule de cocycle multiplicatif (19) est vérifiée pour tous les  $m, n \in \mathbb{Z}$ .

**Remarques.** (1) En munissant  $X \times \mathbb{K}^d$  de la tribu produit, considérons le système dynamique mesurable à temps discret  $(X \times \mathbb{K}^d, \phi_M)$  où

$$\phi_M : (x, v) \mapsto (\phi(x), M(x)v).$$

L'application de première projection  $(x, v) \mapsto x$  de  $X \times \mathbb{K}^d$  dans  $X$  est une semi-conjugaison de  $(X \times \mathbb{K}^d, \phi_M)$  dans  $(X, \phi)$ , et donc  $(X \times \mathbb{K}^d, \phi_M)$  est une extension mesurable de  $(X, \phi)$ . Une récurrence immédiate montre que

$$\forall n \in \mathbb{N} \setminus \{0\}, \quad (\phi_M)^n(x, v) = (\phi^n(x), M_n(x)v),$$

ce qui justifie l'ordre du produit de matrices que nous considérons. Si  $(X, \mathcal{B}, \phi)$  est inversible et si  $M$  est à valeurs dans  $\text{GL}_d(\mathbb{K})$ , alors  $(X \times \mathbb{K}^d, \phi_M)$  est inversible.

(2) Une interprétation probabiliste est la suivante. Soit  $(A_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires indépendantes identiquement distribuées à valeurs dans  $\mathcal{M}_d(\mathbb{K})$  de loi commune  $\rho$ , définies sur un espace de probabilité  $(\tilde{\Omega}, \tilde{\mathcal{B}}, \mathbb{P})$  (voir la note de bas de page 32 pour des définitions). Le *produit aléatoire de matrices* associé à  $(A_n)_{n \in \mathbb{N}}$  est la suite de variables aléatoires  $(A_n A_{n-1} \dots A_2 A_1 A_0)_{n \in \mathbb{N}}$ , qui est aussi à valeurs dans  $\mathcal{M}_d(\mathbb{K})$ . Le lien avec les produits dynamiques de matrices définis ci-dessus est le suivant. Considérons le système dynamique probabilisé  $(X = \mathcal{M}_d(\mathbb{K})^{\mathbb{N}}, \mu = \rho^{\mathbb{N}}, \phi = \sigma_+)$  où  $\phi = \sigma_+$  est le décalage (à gauche) sur l'espace mesurable produit  $\mathcal{M}_d(\mathbb{K})^{\mathbb{N}}$  et  $\mu = \rho^{\mathbb{N}}$  est la mesure de Bernoulli associée à la mesure  $\rho$ . Soit  $M : X \rightarrow \mathcal{M}_d(\mathbb{K})$  l'application mesurable  $(x_i)_{i \in \mathbb{N}} \mapsto x_0$  de première composante. Alors pour tous les  $x = (x_i)_{i \in \mathbb{N}} \in X$  et  $n \in \mathbb{N} \setminus \{0\}$ , nous avons

$$M_n(x) = M(\sigma_+^{n-1}(x)) \dots M(\sigma_+(x)) M(x) = x_{n-1} x_{n-2} \dots x_2 x_1 x_0.$$

Par conséquent, si  $\Theta : \tilde{\Omega} \rightarrow X$  est l'application mesurable  $\omega \mapsto (A_n(\omega))_{n \in \mathbb{N}}$ , qui vérifie  $\Theta_* \mathbb{P} = \mu$  (voir l'exercice E.1) donc préserve la mesure, alors  $M_n \circ \Theta = A_{n-1} \dots A_2 A_1 A_0$ .

(3) Par exemple, si  $X$  est le tore  $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ , et si  $\phi : \mathbb{T}^d \rightarrow \mathbb{T}^d$  est un difféomorphisme  $C^1$ , vu comme un difféomorphisme  $C^1$  de  $\mathbb{R}^d$  qui est  $\mathbb{Z}^d$ -périodique, alors l'application continue  $M : X \rightarrow \text{GL}(\mathbb{R}^d)$  définie par  $x \bmod \mathbb{Z}^d \mapsto d_x \phi$  fournit (en identifiant de manière usuelle  $\text{GL}(\mathbb{R}^d)$  et  $\text{GL}_d(\mathbb{R})$ ) un produit dynamique de matrices pour le système dynamique différentiable  $(X, \phi)$ .

La notion de cocycle multiplicatif peut s'étendre de nombreuses manières, par exemple en une famille  $(M_t)_{t \in \mathbb{R}}$  d'applications de  $X$  dans  $\text{GL}_d(\mathbb{K})$  vérifiant la propriété (19) pour tous les temps  $m, n \in \mathbb{R}$ . Une autre extension<sup>64</sup>, généralisant le cas du fibré tangent trivialisable  $T\mathbb{T}^d = \mathbb{T}^d \times \mathbb{R}^d$  du tore  $\mathbb{T}^d$  précédent est la suivante. Soient  $p : E \rightarrow X$  un fibré vectoriel mesurable au-dessus de  $X$ , et  $\text{End}_\phi(p)$  le fibré vectoriel mesurable sur  $X$  de

64. pour la lectrice ou le lecteur qui connaît les fibrés vectoriels

fibre au-dessus de  $x$  l'espace vectoriel  $\text{End}(E_x, E_{\phi(x)})$  (des applications linéaires de la fibre  $E_x = p^{-1}(x)$  de  $p$  au-dessus de  $x$  dans la fibre  $E_{\phi(x)} = p^{-1}(\phi(x))$  de  $p$  au-dessus de  $\phi(x)$ ). Une section mesurable  $M \in \Gamma(\text{End}_{\phi}(p))$  du fibré  $\text{End}_{\phi}(p)$  (c'est-à-dire une application mesurable  $M : X \rightarrow \text{End}_{\phi}(p)$  telle que  $M(x) \in \text{End}(E_x, E_{\phi(x)})$  pour tout  $x \in X$ ) définit un produit dynamique d'endomorphismes  $(M_n \in \Gamma(\text{End}_{\phi^n}(p)))_{n \in \mathbb{N}}$  par la formule (18) (en remplaçant le produit matriciel par la composition des endomorphismes).

(4) Lorsque  $\| \cdot \|$  est une norme *matricielle*<sup>65</sup> sur l'espace vectoriel  $\mathcal{M}_n(\mathbb{K})$  sur  $\mathbb{K}$ , la suite  $(f_n = \ln \|M_n\|)_{n \in \mathbb{N} \setminus \{0\}}$  est une suite sous-additive de fonctions mesurables réelles et  $f_1^+ = \ln^+ \|M\|$ .

Le résultat suivant (voir [FuK]) est donc une conséquence du théorème de Kingman 2.6 et de l'équivalence de deux normes en dimension finie.

**Corollaire 2.11. (Théorème de Furstenberg<sup>57</sup>-Kesten<sup>57</sup>)** *Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé,  $M : X \rightarrow \mathcal{M}_d(\mathbb{K})$  une application mesurable et  $(M_n)_{n \in \mathbb{N}}$  le produit dynamique de matrice associé. Si  $\ln^+ \|M\|$  est  $\mu$ -intégrable, alors il existe une fonction<sup>66</sup>  $\lambda_{\max} : X \rightarrow \mathbb{R}$  mesurable et  $\mu$ -presque partout  $\phi$ -invariante, telle que  $\mu$ -presque partout, nous ayons la convergence*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n\| = \lambda_{\max} . \quad (20)$$

De plus, la fonction  $\lambda_{\max}^+$  est  $\mu$ -intégrable, et

$$\int_X \lambda_{\max} d\mu = \lim_{n \rightarrow +\infty} \frac{1}{n} \int_X \ln \|M_n\| d\mu = \inf_{n \in \mathbb{N} \setminus \{0\}} \frac{1}{n} \int_X \ln \|M_n\| d\mu \in [-\infty, +\infty[ . \quad \square$$

Soit  $\overline{\mathbb{K}}$  une clôture algébrique<sup>67</sup> de  $\mathbb{K}$ . Rappelons que le *rayon spectral* d'une matrice  $A \in \mathcal{M}_d(\mathbb{K})$  est le nombre réel

$$\rho(A) = \max_{\lambda \in \text{Sp}_{\overline{\mathbb{K}}}(A)} |\lambda| = \lim_{n \rightarrow +\infty} \|A^n\|^{1/n} = \inf_{n \in \mathbb{N} \setminus \{0\}} \|A^n\|^{1/n} . \quad 68$$

Remarquons que si  $M : X \rightarrow \mathcal{M}_d(\mathbb{K})$  est une application constante d'image  $A$  et si  $\phi = \text{id}$ , alors  $M_n = A^n$ , donc  $\ln \|M_n\| \sim n \ln \rho(A)$ , et par conséquent  $\lambda_{\max} = \ln \rho(A)$  est le logarithme du rayon spectral de  $M$ .

Sous les hypothèses générales du théorème de Furstenberg-Kesten 2.11, la fonction obtenue  $\lambda_{\max}$  représente donc le comportement asymptotique presque partout du plus grand module de valeur propre du produit dynamique de matrices  $(M_n)_{n \in \mathbb{N} \setminus \{0\}}$ . Le théorème d'Oseledets 2.12 suivant donne le comportement asymptotique presque partout non seulement des autres valeurs singulières mais aussi des espaces stables associés (voir les définitions ci-dessous).

Faisons quelques rappels d'algèbre linéaire. Soit  $d \in \mathbb{N} \setminus \{0\}$ . Notons  $\text{GL}_d^1(\mathbb{K})$  le sous-groupe de  $\text{GL}_d(\mathbb{K})$  constitué des éléments dont la valeur absolue du déterminant est égale à 1.

65. c'est-à-dire vérifiant  $\|AB\| \leq \|A\| \|B\|$  pour tous les  $A, B \in \mathcal{M}_n(\mathbb{K})$

66. appelée l'*exposant de Lyapounov supérieur* ou *maximal* de  $(X, \mathcal{B}, \mu, \phi, M)$  et aussi notée  $\chi^+$  dans la littérature

67. bien entendu, nous prenons  $\overline{\mathbb{K}} = \mathbb{C}$  si  $\mathbb{K} = \mathbb{R}$  et  $\overline{\mathbb{K}} = \mathbb{K}$  si  $\mathbb{K} = \mathbb{C}$

68. L'existence de la limite et la dernière égalité découlent du lemme de Fekete 2.7. La limite est indépendante de la norme, par leur équivalence en dimension finie. Voir par exemple [Pau4, Théo. 1.25] pour la seconde égalité, où  $\text{Sp}_{\mathbb{C}}(A)$  est l'ensemble des valeurs propres complexes de  $A$ .

**Valeurs singulières.** Nous renvoyons par exemple à [Pau7, §1.10] pour les rappels suivants sur la *décomposition en valeurs singulières* (SVD en anglais). Pour toute matrice  $A \in \mathcal{M}_d(\mathbb{K})$ , en notant  $A^* = {}^t\bar{A}$  sa matrice adjointe<sup>69</sup>, la matrice produit  $A^*A$  est symétrique et positive, donc diagonalisable en base orthonormée de valeurs propres positives ou nulles. Les *valeurs singulières*

$$(\sigma_1(A), \dots, \sigma_d(A))$$

de  $A$  sont les racines carrées des valeurs propres de  $A^*A$  ordonnées dans l'ordre décroissant :  $\sigma_1(A) \geq \dots \geq \sigma_d(A) \geq 0$  (et celles nulles sont parfois omises dans la littérature). Les *espaces singuliers* (à droite) de  $A$  sont les espaces propres de  $A^*A$ , et les *espaces singuliers* (à gauche) de  $A$  sont les espaces propres de  $AA^*$ . L'application  $A \mapsto (\sigma_1(A), \dots, \sigma_d(A))$  de  $\mathcal{M}_d(\mathbb{K})$  dans  $[0, +\infty[^d$  est continue. De plus (voir loc. cit.), la matrice  $A^*$  a les mêmes valeurs singulières que  $A$ , et il existe une matrice  $D \in \mathcal{M}_d(\mathbb{R})$  diagonale, dont les coefficients diagonaux sont (dans l'ordre) les valeurs singulières de  $A$ , et des matrices orthogonales  $P, P' \in O(d)$  si  $\mathbb{K} = \mathbb{R}$  et unitaires  $P, P' \in U(d)$  si  $\mathbb{K} = \mathbb{C}$  telles que<sup>70</sup>

$$A = P' D P^{-1} .$$

Prenons  $\| \cdot \|$  la norme usuelle sur  $\mathbb{K}^d$ , et  $\| \cdot \|$  la norme d'opérateur associée sur  $\mathcal{M}_d(\mathbb{K})$ . Alors puisque les éléments des groupes orthogonaux ou unitaires préservent la norme d'opérateur par multiplication à droite ou à gauche, nous avons

$$\|A\| = \|D\| = \sigma_1(A) .$$

De plus, si  $A \in \text{GL}_d(\mathbb{K})$ , alors les valeurs singulières sont strictement positives, et nous avons  $A^{-1} = P D^{-1} P'^{-1}$ , donc en réordonnant les coefficients diagonaux de  $D^{-1}$ , nous avons

$$\sigma_1(A^{-1}) = \sigma_d(A)^{-1}, \dots, \sigma_d(A^{-1}) = \sigma_1(A)^{-1} .$$

En particulier

$$\sigma_d(A) = \|A^{-1}\|^{-1} .$$

Si  $A \in \text{GL}_d^1(\mathbb{K})$ , alors  $A^*A \in \text{SL}_d(\mathbb{K})$  car  $\det(A^*A) = |\det A|^2$ , donc

$$\prod_{i=1}^d \sigma_i(A) = 1 .$$

De plus, si  $(e_1, \dots, e_d)$  est la base canonique de  $\mathbb{K}^d$ , alors  $(Pe_1, \dots, Pe_d)$  est une base orthonormée de  $\mathbb{K}^d$  dans laquelle la matrice de  $A^*A = PD^2P^{-1}$  est diagonale, et

$$\forall i \in \llbracket 1, d \rrbracket, \quad \|APe_i\| = \|P'De_i\| = \sigma_i(A)\|P'e_i\| = \sigma_i(A) .$$

En particulier  $\|APe_1\| = \|A\|$  et  $\|APe_d\| = \|A^{-1}\|^{-1}$ . Remarquons encore que les vecteurs  $APe_1 = \sigma_1 P'e_1, \dots, APe_d = \sigma_d P'e_d$  sont deux à deux orthogonaux car la matrice  $P'$  est orthogonale.

69. Pour tout  $A = (a_{ij})_{1 \leq i, j \leq d} \in \mathcal{M}_d(\mathbb{K})$ , nous avons  $A^* = (\bar{a}_{ji})_{1 \leq i, j \leq d}$ . Notons que comme  $\bar{\bar{x}} = x$  pour tout  $x \in \mathbb{R}$ , nous avons  $A^* = {}^t A$  si  $A \in \mathcal{M}_d(\mathbb{R})$ .

70. Cette écriture est appelée la *décomposition de Cartan* de  $A$ , nous y reviendrons plus longement dans le lemme 9.10. L'application  $A \mapsto (\sigma_1(A), \dots, \sigma_d(A))$  de  $\text{GL}_d(\mathbb{K})$  dans  $[0, +\infty[^d$  est appelée la *projection de Cartan* de  $\text{GL}_d(\mathbb{K})$ .

**Espaces de drapeaux.** Comme les espaces singuliers (à droite) des produits aléatoires de matrices vont dépendre du point de l'espace des phases considéré, il est important de comprendre la structure mesurable des ensembles de sous-espaces vectoriels.

Soit  $k \in \llbracket 1, d \rrbracket$ . La *variété grassmannienne*  $\text{Gr}_{d,k} = \text{Gr}_{d,k}(\mathbb{K})$  de rang  $k$  en dimension  $d$  sur  $\mathbb{K}$  est l'espace topologique<sup>71</sup> (métrisable séparable) des sous-espaces vectoriels de dimension  $k$  de  $\mathbb{K}^d$ . Un *drapeau* de  $\mathbb{K}^d$  de longueur  $k$  est une suite  $(V_i)_{1 \leq i \leq k}$  strictement décroissante de sous-espaces vectoriels non nuls de  $\mathbb{K}^d$ . Par convention, nous posons  $V_{k+1} = \{0\}$  si  $(V_i)_{1 \leq i \leq k}$  est un drapeau de longueur  $k$ . La *dimension* d'un drapeau  $(V_i)_{1 \leq i \leq k}$  est le  $k$ -uplet strictement décroissant d'entiers strictement positifs  $(\dim V_1, \dots, \dim V_k)$ . Si  $(d_1, \dots, d_k)$  est un  $k$ -uplet strictement décroissant d'entiers strictement positifs, l'*espace des drapeaux de dimension*  $(d_1, \dots, d_k)$  de  $\mathbb{K}^d$  est le sous-espace topologique (métrisable séparable) de l'espace topologique produit  $\text{Gr}_{d,d_1} \times \dots \times \text{Gr}_{d,d_k}$  défini par

$$\text{Drap}_{(d_1, \dots, d_k)} = \{(V_1, \dots, V_k) \in \text{Gr}_{d,d_1} \times \dots \times \text{Gr}_{d,d_k} : V_1 \supset \dots \supset V_k\}.$$

L'*espace des drapeaux* de  $\mathbb{K}^d$  est l'espace topologique (métrisable séparable) somme disjointe

$$\text{Drap}(\mathbb{K}^d) = \bigsqcup_{k, d_1, \dots, d_k \in \llbracket 1, d \rrbracket : d_1 > \dots > d_k} \text{Drap}_{(d_1, \dots, d_k)}.$$

Comme convenu, tout espace topologique est muni de sa tribu borélienne.

**Théorème 2.12. (Théorème ergodique multiplicatif d'Oseledets)** *Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret. Soit  $M : X \rightarrow \text{GL}_d(\mathbb{R})$  une application mesurable telle que les deux applications  $\ln^+ \|M\|$  et  $\ln^+ \|M^{-1}\|$  soient  $\mu$ -intégrables. Notons  $(M_n)_{n \in \mathbb{N}}$  le produit dynamique matriciel à gauche associé à  $(\phi, M)$ .*

*Alors il existe une partie mesurable  $\phi$ -invariante  $X'$  de mesure pleine dans  $X$ , une application mesurable  $V : X' \rightarrow \text{Drap}(\mathbb{K}^d)$ , et des applications mesurables  $\phi$ -invariantes  $k : X' \rightarrow \llbracket 1, d \rrbracket$  et  $\lambda : X' \mapsto \bigsqcup_{i \in \llbracket 1, d \rrbracket} \mathbb{R}^i$ , qui sont presque partout uniques, telles que pour tout  $x \in X'$ ,*

$$\lambda(x) = (\lambda_1(x), \dots, \lambda_{k(x)}(x)) \quad \text{et} \quad \lambda_1(x) > \dots > \lambda_{k(x)}(x),$$

$$V(x) = (V_1(x), \dots, V_{k(x)}(x)) \quad \text{et} \quad \forall i \in \llbracket 1, k(x) \rrbracket, \quad M(x)V_i(x) = V_i(\phi(x)),$$

$$\forall i \in \llbracket 1, k(x) \rrbracket, \forall v \in V_i(x) \setminus V_{i+1}(x), \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| = \lambda_i(x),$$

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \ln |\det(M_n(x))| = \sum_{i=1}^{k(x)} \lambda_i(x) (\dim V_i(x) - \dim V_{i+1}(x)). \quad (21)$$

*Si  $(X, \mathcal{B}, \mu, \phi)$  est ergodique, alors les fonctions  $k, \lambda_1, \dots, \lambda_k$  sont presque partout constantes.*

71. C'est l'espace topologique quotient du sous-espace topologique de  $(\mathbb{K}^d)^k$  constitué des  $k$ -uplets de vecteurs linéairement indépendants par la relation d'équivalence « engendrer le même sous-espace vectoriel » : deux sous-espaces vectoriels de même dimension sont proches s'ils admettent des bases proches formées de vecteurs de norme 1. La variété grassmannienne  $\text{Gr}_{d,k}$  est ainsi métrisable en définissant, pour tous les  $V, V' \in \text{Gr}_{d,k}$ , la distance  $d(V, V') = \min\{\sup_{1 \leq i \leq k} \|e_i - e'_i\|\}$  où la borne inférieure (atteinte par compacité) est prise sur toutes les bases  $(e_1, \dots, e_k)$  de  $V$  et  $(e'_1, \dots, e'_k)$  de  $V'$  constituées de vecteurs de norme 1.

Si de plus  $(X, \mathcal{B}, \mu, \phi)$  est inversible, alors il existe une application mesurable notée  $E : x \mapsto (E_1(x), \dots, E_{k(x)}(x))$  de  $X$  à valeurs dans l'espace topologique somme disjointe  $\bigsqcup_{k, \delta_1, \dots, \delta_k \in \llbracket 1, d \rrbracket} \text{Gr}_{d, \delta_1} \times \dots \times \text{Gr}_{d, \delta_k}$  telle que pour tout  $x \in X'$ , nous ayons<sup>72</sup>

$$\forall i \in \llbracket 1, k(x) \rrbracket, \quad V_i(x) = V_{i+1}(x) \oplus E_i(x), \quad M(x)E_i(x) = E_i(\phi(x)),$$

$$\forall v \in E_i(x) \setminus \{0\}, \quad \lim_{n \rightarrow -\infty} \frac{1}{n} \ln \|M_n(x)v\| = \lambda_i(x).$$

Pour tout  $k' \in \llbracket 1, d \rrbracket$ , les applications<sup>73</sup>  $\lambda_1, \dots, \lambda_{k'}$  sont appelées les *fonctions de Lyapounov* du produit dynamique de matrices  $(M_n)_{n \in \mathbb{N}}$  sur  $\{x \in X' : k(x) = k'\}$ . Pour tout  $x \in X'$ , les nombres réels  $\lambda_1(x), \dots, \lambda_{k(x)}(x)$  (respectivement les sous-espaces vectoriels  $V_1(x), \dots, V_{k(x)}(x)$ ) sont appelés les *exposants caractéristiques* ou *exposants de Lyapounov* (respectivement les *sous-espaces de Lyapounov*) en  $x$  du produit dynamique de matrices  $(M_n)_{n \in \mathbb{N}}$ . L'ensemble  $\{\lambda_1(x), \dots, \lambda_{k(x)}(x)\}$  est appelé le *spectre de Lyapounov* en  $x$  du produit dynamique de matrices  $(M_n)_{n \in \mathbb{N}}$ .

Par exemple, si  $\phi = \text{id}$  (inversible) et  $M : X \rightarrow \text{GL}_d(\mathbb{K})$  est une application constante, de valeur  $A \in \text{GL}_d(\mathbb{K})$ , alors par les rappels ci-dessus sur les valeurs singulières de  $A$ , nous avons

- l'application  $M_n$  est constante de valeur  $A^n$ ,
- l'application  $k$  est constante, égale au nombre de valeurs singulières deux à deux distinctes de  $A$ ,
- les applications  $\lambda_1, \dots, \lambda_k$  sont constantes, égales aux logarithmes des valeurs singulières deux à deux distinctes de  $A$  ordonnées de manière strictement décroissante,
- les applications  $E_1, \dots, E_k$  sont constantes, égales aux espaces propres pour les valeurs propres  $e^{2\lambda_1}, \dots, e^{2\lambda_k}$  de  $A^*A$ .

**Remarques.** (1) Le théorème de Furstenberg-Kesten 2.11 (appliqué à l'action de  $M$  sur la  $k$ -ème puissance extérieure  $\Lambda^k(\mathbb{K}^d)$  de  $\mathbb{K}^d$  pour  $k \in \llbracket 1, d \rrbracket$ ) donne en toute dimension une expression des exposants de Lyapounov, comme suit (voir par exemple [Arn, §3]). Pour tout  $i \in \llbracket 1, d \rrbracket$  et  $\mu$ -presque tout  $x \in X$ , les limites

$$\widehat{\lambda}_i(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \sigma_i(M_n(x))$$

existent. Nous avons  $\widehat{\lambda}_1(x) \geq \dots \geq \widehat{\lambda}_d(x) \geq 0$ . Alors  $k(x) \in \llbracket 1, d \rrbracket$  est le nombre de ces limites qui sont deux à deux distinctes, et  $\lambda_1(x), \dots, \lambda_{k(x)}(x)$  sont celles deux à deux distinctes, rangées dans l'ordre strictement décroissant.

(2) Précisons encore plus comment obtenir en pratique les exposants de Lyapounov et les sous-espaces de Lyapounov, après quelques rappels sur l'application exponentielle des matrices.

72. Bien entendu, pour tous les  $x \in X'$  et  $i \in \llbracket 1, k(x) \rrbracket$ , en posant  $\delta_i(x) = \dim E_i(x)$ , nous avons alors

$$\delta_i(x) = \dim V_i(x) - \dim V_{i+1}(x), \quad V_i(x) = \bigoplus_{j=i}^{k(x)} E_j(x),$$

$$\forall v \in E_i(x) \setminus \{0\}, \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| = \lambda_i(x), \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \ln |\det(M_n(x))| = \sum_{i=1}^{k(x)} \lambda_i(x) \delta_i(x).$$

73. On note parfois  $\lambda_{\max} = \chi^+ = \lambda_1$ , et  $\lambda_{\min} = \chi^- = \lambda_k$ , appelé l'*exposant de Lyapounov inférieur* ou *minimal*.

Notons  $\text{Herm}_d = \{A \in \mathcal{M}_d(\mathbb{K}) : A^* = 1\}$  l'espace vectoriel réel de dimension finie des matrices hermitiennes (appelées symétriques si  $\mathbb{K} = \mathbb{R}$ , et on note alors cet espace vectoriel  $\text{Sym}_d$ !) de  $\mathcal{M}_d(\mathbb{K})$ , et  $\text{Herm}_d^{++}$  son ouvert formé des éléments définis positifs (noté  $\text{Sym}_d^{++}$  si  $\mathbb{K} = \mathbb{R}$ ). Rappelons (voir par exemple [Pau7]) que l'application exponentielle des matrices  $\exp$  est un  $C^\infty$ -difféomorphisme de  $\text{Herm}_d$  dans  $\text{Herm}_d^{++}$ . Pour tous les  $A \in \text{Herm}_d^{++}$  et  $\alpha \in \mathbb{R}$ , posons alors

$$A^\alpha = \exp(\alpha \exp^{-1}(A))$$

et remarquons que l'application  $A \mapsto A^\alpha$  est un  $C^\infty$ -difféomorphisme de  $\text{Herm}_d^{++}$  dans lui-même.

Avec les notations (et les hypothèses) du théorème d'Oseledets 2.12, il est possible de montrer (voir par exemple [Arn, §3]) que pour  $\mu$ -presque tout  $x \in X$ , la limite

$$A(x) = \lim_{n \rightarrow +\infty} (M_n(x)^* M_n(x))^{\frac{1}{2n}}$$

existe, que c'est une matrice hermitienne définie positive (ceci utilise la seconde hypothèse d'intégrabilité de  $\ln \|M^{-1}\|$ ), et que si  $a_1(x) > a_2(x) > \dots > a_{k(x)}(x)$  sont les valeurs propres distinctes de  $A(x)$ , rangées dans l'ordre strictement décroissant, d'espaces propres associés  $E_1(x), E_2(x), \dots, E_{k(x)}(x)$ , alors pour tout  $i \in \llbracket 1, k(x) \rrbracket$ , nous avons

$$\lambda_i(x) = \ln a_i(x) \quad \text{et} \quad V_i(x) = \bigoplus_{j=i}^{k(x)} E_j(x).$$

Par la continuité du déterminant et du logarithme, nous avons

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{1}{n} \ln |\det(M_n(x))| &= \lim_{n \rightarrow +\infty} \frac{1}{2n} \ln \det(M_n(x)^* M_n(x)) \\ &= \lim_{n \rightarrow +\infty} \ln \det((M_n(x)^* M_n(x))^{\frac{1}{2n}}) = \ln \det A(x) = \ln \left( \prod_{i=1}^{k(x)} a_i(x)^{\dim E_i(x)} \right) \\ &= \sum_{i=1}^{k(x)} \lambda_i(x) \dim E_i(x), \end{aligned}$$

ce qui redémontre la formule (21).

**Démonstration.** <sup>74</sup> Soient  $X, \mathcal{B}, \mu, \phi, M, (M_n)_{n \in \mathbb{N}}$  comme dans l'énoncé. Par l'équivalence des normes en dimension finie, nous pouvons supposer que la norme  $\| \cdot \|$  sur  $\mathcal{M}_d(\mathbb{K})$  est la norme d'opérateur pour la norme usuelle  $\| \cdot \|$  sur  $\mathbb{K}^d$ .

**Étape 1.** Montrons que nous pouvons supposer que l'application  $M : X \rightarrow \text{GL}_d(\mathbb{K})$  soit à valeurs dans  $\text{GL}_d^1(\mathbb{K})$ .

Considérons les applications mesurables  $f = \ln |\det M| : X \rightarrow \mathbb{R}$ , ses moyennes de Birkhoff  $S_n f = \frac{1}{n} \sum_{k=0}^{n-1} f \circ \phi^k : X \rightarrow \mathbb{R}$  et  $M' = e^{-\frac{1}{d} S_n f} M : X \rightarrow \text{GL}_d(\mathbb{K})$ . Remarquons que  $|\det M'| = e^{-f} |\det M| = 1$ , donc  $M'$  est à valeurs dans  $\text{GL}_d^1(\mathbb{K})$ . Notons  $(M'_n)_{n \in \mathbb{N}}$  le produit dynamique matriciel à gauche associé à  $(\phi, M')$ . Pour tout  $n \in \mathbb{N} \setminus \{0\}$ , puisque

<sup>74</sup>. Elle est spécifique aux cas  $\mathbb{K} = \mathbb{R}$  et  $\mathbb{K} = \mathbb{C}$ , et doit être contournée pour les corps commutatifs valués localement compacts.



les homothéties sont centrales dans  $\mathrm{GL}_d(\mathbb{K})$ , nous avons  $M'_n = e^{-nS_n f} M_n$ . Donc pour tout  $v \in \mathbb{K}^d \setminus \{0\}$ , nous avons

$$\frac{1}{n} \ln \|M_n(x)v\| = S_n f(x) + \frac{1}{n} \ln \|M'_n(x)v\| .$$

Pour toute matrice  $A \in \mathcal{M}_d(\mathbb{K})$ , nous avons

$$|\det A| = \prod_{\lambda \in \mathrm{Sp}_{\mathbb{K}} A} |\lambda| \leq \left( \sup_{\lambda \in \mathrm{Sp}_{\mathbb{K}} A} |\lambda| \right)^d = \rho(A)^d = \left( \inf_{n \in \mathbb{N}} \|A^n\|^{\frac{1}{n}} \right)^d \leq \|A\|^d .$$

L'hypothèse que  $\ln^+ \|M\|$  est intégrable implique donc que  $f^+ = \ln^+ |\det M| \leq d \ln^+ \|M\|$  est intégrable. De même, l'hypothèse que  $\ln^+ \|M^{-1}\|$  est intégrable implique que  $(-f)^+ = \ln^+ |\det M|^{-1} = \ln^+ |\det(M^{-1})| \leq d \ln^+ \|M^{-1}\|$  est intégrable. Donc  $f$  est intégrable. Par le théorème de Birkhoff 2.2, la suite  $(S_n f)_{n \in \mathbb{N}}$  converge presque partout vers une application mesurable  $\tilde{f} : X \rightarrow \mathbb{R}$ . Donc si  $\lambda'_1, \dots, \lambda'_{k'}, V'_1, \dots, V'_{k'}$  sont les exposants de Lyapounov et sous-espaces de Lyapounov de  $(\phi, M')$  supposés exister, alors en posant (en dehors d'un ensemble de mesure nulle)  $k = k'$ , et pour tout  $i \in \llbracket 1, k \rrbracket$ ,

$$\lambda_i = \tilde{f} + \lambda'_i \quad \text{et} \quad V_i = V'_i ,$$

ainsi que  $E_i = E'_i$  dans le cas inversible, le théorème d'Oseledets est valable pour  $M$ .

Nous n'allons maintenant donner la démonstration du théorème d'Oseledets que lorsque  $d = 2$  et sans le complément obtenu en supposant que  $\phi$  est inversible. Nous renvoyons par exemple à [Via, KM], ainsi qu'à [Arn, §3] (cette dernière référence donnant explicitement l'extension au cas du temps continu), pour une démonstration complète. Le cas  $d = 2$  contient toutefois l'essentiel des idées de la démonstration.

Appliquons le théorème de Furstenberg-Kesten 2.11 d'une part à  $M$  et d'autre part à  $M^{-1}$  en renversant l'ordre des matrices, le résultat restant valable. Ceci est possible par les deux hypothèses d'intégrabilité du théorème d'Oseledets 2.12. Alors, quitte à enlever un ensemble de mesure nulle de  $X$ , nous pouvons supposer que les limites

$$\lambda_{\max} = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n\| \quad \text{et} \quad \lambda_{\min} = - \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n^{-1}\| \quad (22)$$

existent partout, et sont des fonctions mesurables et  $\phi$ -invariantes. En fait, comme  $\|A\| = \|A^{-1}\|$  pour toute matrice  $A \in \mathrm{GL}_2^1(\mathbb{R})$  (car le produit des deux valeurs singulières de  $A$  vaut 1, voir les rappels ci-dessus), nous avons que  $\lambda_{\max}(x)$  existe si et seulement si  $\lambda_{\min}(x)$  existe, et alors

$$\lambda_{\min}(x) = -\lambda_{\max}(x) .$$

De plus, comme  $\|A\|^2 \geq |\det A| = 1$ , nous avons  $\lambda_{\max}(x) \geq 0$  et donc  $\lambda_{\min}(x) \leq 0$ .

**Étape 2.** (Cas d'égalité des deux exposants de Lyapounov) Considérons l'ensemble mesurable  $\phi$ -invariant  $X_{=0} = \{y \in X : \lambda_{\max}(y) = 0\}$ . Montrons que pour tous les  $x \in X_{=0}$ , nous avons

$$\forall v \in \mathbb{K}^2 \setminus \{0\}, \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| = \lambda_{\max}(x) .$$

Il suffira alors pour tout  $x \in X_{=0}$  de poser  $k(x) = 1$ ,  $\lambda_1(x) = \lambda_{\max}(x)$  et  $V_1(x) = \mathbb{K}^2$ .

Si  $x \in X_{=0}$ , alors  $\lambda_{\max}(x) = \lambda_{\min}(x) = 0$ . Soit  $v \in \mathbb{K}^2 \setminus \{0\}$ . Par la définition de la norme d'opérateur, nous avons

$$\|M_n(x)^{-1}\|^{-1}\|v\| \leq \|M_n(x)v\| \leq \|M_n(x)\| \|v\| .$$

En prenant le logarithme, et en divisant par  $n$ , les membres de gauche et de droite de ces inégalités tendent vers 0, ce qui montre que  $\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| = 0$ , comme voulu.

**Étape 3.** (Cas d'inégalité des deux exposants de Lyapounov) Considérons l'ensemble mesurable  $\phi$ -invariant  $X_{>0} = \{y \in X : \lambda_{\max}(y) > 0\}$ . Montrons qu'il existe une application mesurable  $E^s : X_{>0} \rightarrow \text{Gr}_{1,2}$  telle que pour presque tout  $x \in X_{>0}$ , nous ayons la propriété d'équivariance  $M(x)E^s(x) = E^s(\phi(x))$  et

$$\forall v \in \mathbb{K}^2 \setminus E^s(x), \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| = \lambda_{\max}(x) , \quad (23)$$

$$\forall v \in E^s(x) \setminus \{0\}, \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| = \lambda_{\min}(x) . \quad (24)$$

Il suffira alors pour presque tout élément  $x \in X_{>0}$  de poser  $k(x) = 2$ ,  $\lambda_1(x) = \lambda_{\max}(x)$ ,  $\lambda_2(x) = \lambda_{\min}(x)$ ,  $V_1(x) = \mathbb{K}^2$  et  $V_2(x) = E^s(x)$ .

Soit  $x \in X_{>0}$ . Nous avons en particulier  $\lambda_{\max}(x) > \lambda_{\min}(x)$ . Pour tout  $n \in \mathbb{N} \setminus \{0\}$  assez grand, nous avons donc  $\|M_n(x)\| > \|M_n(x)^{-1}\|^{-1}$ , et les deux valeurs singulières  $\sigma_1(M_n(x)) = \|M_n(x)\|$  et  $\sigma_2(M_n(x)) = \|M_n(x)^{-1}\|^{-1}$  de  $M_n(x)$  sont distinctes. Notons alors  $(u_n(x), s_n(x))$  une base orthonormée de  $\mathbb{K}^2$  dans laquelle la matrice  $M_n(x)^*M_n(x)$  est diagonale, telle que

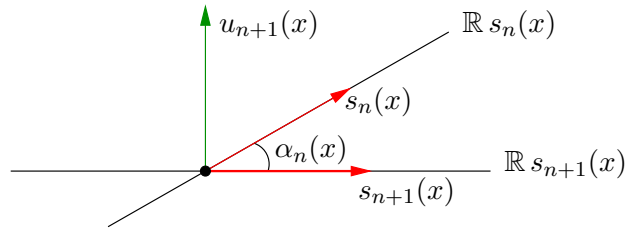
$$M_n(x)u_n(x) = \|M_n(x)\|u_n(x) \quad \text{et} \quad M_n(x)s_n(x) = \|M_n(x)^{-1}\|^{-1}s_n(x)$$

(voir les rappels ci-dessus sur les valeurs singulières). Donc

$$\|M_n(x)u_n(x)\| = \|M_n(x)\| \quad \text{et} \quad \|M_n(x)s_n(x)\| = \|M_n(x)^{-1}\|^{-1} , \quad (25)$$

et les vecteurs  $M_n(x)u_n(x)$  et  $M_n(x)s_n(x)$  sont orthogonaux. Notons  $\alpha_n(x) \in [0, \pi[$  l'angle entre les droites  $\mathbb{K}s_n(x)$  et  $\mathbb{K}s_{n+1}(x)$ . Comme les vecteurs  $u_n(x)$  et  $s_n(x)$  sont uniques modulo multiplication par un élément de  $\mathbb{K}$  de module 1, nous pouvons supposer par récurrence que  $\alpha_n(x)$  est l'angle entre les vecteurs  $s_n(x)$  et  $s_{n+1}(x)$ , et que

$$s_n(x) = \cos(\alpha_n(x))s_{n+1}(x) + \sin(\alpha_n(x))u_{n+1}(x) .$$



L'idée clef du restant de la démonstration du théorème d'Oseledets est de montrer que pour presque tout  $x \in X_{>0}$ , cet angle  $\alpha_n(x)$  tend exponentiellement vite vers 0, de sorte

que la suite de vecteurs unitaires  $(s_n(x))_{n \in \mathbb{N}}$  soit une suite de Cauchy dans  $\mathbb{K}^2$ , et de noter  $E^s(x)$  la droite engendrée par sa limite  $\lim_{n \rightarrow +\infty} s_n(x)$ .

**Affirmation (i).** Pour  $\mu$ -presque tout  $x \in X_{>0}$ , nous avons

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \ln \sin(\alpha_n(x)) \leq -2 \lambda_{\max}(x) .$$

**Démonstration.** Puisque les vecteurs  $M_{n+1}(x)u_{n+1}(x)$  et  $M_{n+1}(x)s_{n+1}(x)$  sont orthogonaux, nous avons par le théorème de Pythagore et la formule (25) de gauche

$$\|M_{n+1}(x)s_n(x)\| \geq \|\sin(\alpha_n(x))M_{n+1}(x)u_{n+1}(x)\| = \sin(\alpha_n(x))\|M_{n+1}(x)\| .$$

Par la définition de la norme d'opérateur et la formule (25) de droite, nous avons

$$\begin{aligned} \|M_{n+1}(x)s_n(x)\| &= \|M(\phi^n(x))M_n(x)s_n(x)\| \leq \|M(\phi^n(x))\| \|M_n(x)s_n(x)\| \\ &= \|M(\phi^n(x))\| \|M_n(x)^{-1}\|^{-1} . \end{aligned}$$

Donc  $0 \leq \sin(\alpha_n(x)) \leq \frac{\|M(\phi^n(x))\|}{\|M_n(x)^{-1}\| \|M_{n+1}(x)\|}$ . En prenant le logarithme, en divisant par  $n$ , et en prenant la limite supérieure quand  $n$  vers  $+\infty$ , l'affirmation (i) découle alors pour presque tout  $x \in X_{>0}$  des définitions (22) de  $\lambda_{\min}$  et  $\lambda_{\max}$ , du fait que  $\lambda_{\min} = -\lambda_{\max}$ , et du lemme 2.13 ci-dessous appliqué à  $f = \ln \|M\|$ , qui est intégrable comme vu dans l'étape 1.  $\square$

**Lemme 2.13.** Soit  $f \in \mathbb{L}^1(X, \mu)$ . Pour  $\mu$ -presque tout  $x \in X$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} f \circ \phi^n(x) = 0 .$$

**Démonstration.** Voir le lemme 2.9 pour une démonstration différente (n'utilisant pas le théorème de Birkhoff). Soit  $g = f \circ \phi - f$ , qui est intégrable. Par le théorème de Birkhoff 2.2, il existe une fonction intégrable  $\phi$ -invariante  $\tilde{g}$  sur  $X$  et une partie mesurable  $X'$  de mesure pleine dans  $X$  telle que pour tout  $x \in X'$ , nous avons  $\lim_{n \rightarrow +\infty} S_n g = \tilde{g}$ . La partie  $A = \{x \in X' : \tilde{g} \neq 0\}$  est mesurable et  $\phi$ -invariante. Nous avons par somme télescopique  $\frac{1}{n} f \circ \phi^n = \frac{1}{n} f + S_n g$ . Donc pour tout  $x \in A$ , nous avons  $\lim_{n \rightarrow +\infty} |f \circ \phi^n(x)| = +\infty$ . Or le lemme de Fatou, et l'invariance de  $A$  et de  $\mu$ , impliquent que

$$\int_A \liminf_{n \rightarrow +\infty} |f \circ \phi^n| d\mu \leq \liminf_{n \rightarrow +\infty} \int_A |f \circ \phi^n| d\mu = \int_A |f| d\mu < +\infty ,$$

donc  $\mu(A) = 0$ , ce qui montre le résultat.  $\square$

Il découle de l'affirmation (i) que pour presque tout point  $x \in X_{>0}$  la suite de vecteurs unitaires  $(s_n(x))_{n \in \mathbb{N}}$  est une suite de Cauchy dans  $\mathbb{K}^2$ , donc converge vers un vecteur unitaire  $s(x) \in \mathbb{K}^2$ . Définissons alors

$$E^s(x) = \mathbb{K} s(x) ,$$

qui dépend de manière mesurable de (presque tout)  $x$ . Remarquons que par les propriétés des suites de Cauchy et des restes de séries géométriques, si  $\beta_n(x) \in [0, \frac{\pi}{2}]$  est l'angle entre  $s_n(x)$  et  $s(x)$  (bien défini pour  $n$  assez grand), alors nous avons

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \ln \sin \beta_n(x) \leq -2 \lambda_{\max}(x) . \quad (26)$$

**Affirmation (ii).** Pour  $\mu$ -presque tout  $x \in X_{>0}$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)s(x)\| = \lambda_{\min}(x),$$

ce qui montre la formule (24).

**Démonstration.** Puisque le minimum de l'application  $v \mapsto \|M_n(x)v\|$  définie sur les vecteurs unitaires de  $\mathbb{K}^2$  est atteint en  $v = s_n(x)$  (voir les rappels ci-dessus sur les valeurs singulières), par la formule (25) de droite et par la définition (22) de  $\lambda_{\min}(x)$ , pour  $\mu$ -presque tout  $x \in X_{>0}$ , nous avons la minoration

$$\begin{aligned} \liminf_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)s(x)\| &\geq \liminf_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)s_n(x)\| \\ &= \liminf_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)^{-1}\|^{-1} = \lambda_{\min}(x). \end{aligned}$$

Montrons que pour  $\mu$ -presque tout  $x \in X_{>0}$ , nous avons aussi la majoration

$$\limsup_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)s(x)\| \leq \lambda_{\min}(x).$$

L'affirmation (ii) en découlera.

Pour  $\mu$ -presque tout  $x \in X_{>0}$ , si  $n$  est assez grand, par la définition de  $\beta_n$ , il existe  $a \in \mathbb{K}$  de valeur absolue 1 tel que

$$s(x) = \cos(\beta_n(x))s_n(x) + a \sin(\beta_n(x))u_n(x).$$

Donc par l'inégalité triangulaire et la formule (25), nous avons

$$\begin{aligned} \|M_n(x)s(x)\| &\leq \cos(\beta_n(x))\|M_n(x)s_n(x)\| + \sin(\beta_n(x))\|M_n(x)u_n(x)\| \\ &= \cos(\beta_n(x))\|M_n(x)^{-1}\|^{-1} + \sin(\beta_n(x))\|M_n(x)\| \\ &\leq 2 \max\{\|M_n(x)^{-1}\|^{-1}, \sin(\beta_n(x))\|M_n(x)\|\}. \end{aligned}$$

En prenant le logarithme, en divisant par  $n$  et en prenant la limite supérieure quand  $n \rightarrow +\infty$ , par la formule (26) et puisque  $\lambda_{\max} = -\lambda_{\min}$  presque partout, pour  $\mu$ -presque tout  $x \in X_{>0}$ , nous avons

$$\begin{aligned} &\limsup_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)s(x)\| \\ &\leq \max\left\{-\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)^{-1}\|, \limsup_{n \rightarrow +\infty} \frac{1}{n} \ln \sin \beta_n(x) + \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)\|\right\} \\ &\leq \max\{\lambda_{\min}(x), -2\lambda_{\max}(x) + \lambda_{\max}(x)\} = \lambda_{\min}(x), \end{aligned}$$

ce qu'il fallait démontrer. □

**Affirmation (iii).** Pour  $\mu$ -presque tout  $x \in X_{>0}$ , nous avons

$$\forall v \in \mathbb{K}^2 \setminus E^s(x), \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| = \lambda_{\max}(x),$$

ce qui montre la formule (23).

**Démonstration.** La démonstration est semblable à celle de l'affirmation (ii). Si  $v \in \mathbb{K}^2$  n'est pas colinéaire à  $s(x)$ , notons  $\gamma_n \in [0, \pi]$  l'angle entre  $v$  et  $s_n(x)$ , qui converge donc vers un angle non nul modulo  $\pi$  quand  $n$  tend vers  $+\infty$ . Alors il existe  $b \in \mathbb{K}$  de valeur absolue 1 tel que, en utilisant la formule (25) de gauche, nous ayons

$$\begin{aligned} \|M_n(x)v\| &= \|M_n(x)((\cos \gamma_n)s_n(x) + b(\sin \gamma_n)u_n(x))\| \geq \sin \gamma_n \|M_n(x)u_n(x)\| \\ &= \sin \gamma_n \|M_n(x)\|. \end{aligned}$$

En prenant le logarithme, en divisant par  $n$  et en prenant la limite inférieure, par la définition (22) de  $\lambda_{\max}(x)$ , nous avons donc  $\liminf_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| \geq \lambda_{\max}(x)$ .

Puisque le maximum de l'application  $w \mapsto \|M_n(x)w\|$  définie sur les vecteurs unitaires de  $\mathbb{K}^2$  est atteint en  $w = u_n(x)$  (voir les rappels ci-dessus sur les valeurs singulières), par la formule (25) de gauche et par la définition (22) de  $\lambda_{\max}(x)$ , pour  $\mu$ -presque tout  $x \in X_{>0}$ , nous avons la majoration

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| &\leq \liminf_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)u_n(x)\| \\ &= \liminf_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)\| = \lambda_{\max}(x). \end{aligned}$$

L'affirmation (iii) en découle. □

**Affirmation (iv).** Pour  $\mu$ -presque tout  $x \in X_{>0}$ , nous avons la propriété d'équivariance  $M(x)E^s(x) = E^s(\phi(x))$ .

**Démonstration.** Pour  $\mu$ -presque tout  $x \in X_{>0}$ , nous avons

$$M_n(\phi(x))M(x)s(x) = M_{n+1}(x)s(x). \quad (27)$$

En regardant le membre de gauche de cette égalité, pour presque tout  $x \in X_{>0}$ , si par l'absurde  $M(x)s(x)$  n'était pas colinéaire à  $s(\phi(x))$ , alors par l'affirmation (iii), nous aurions

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(\phi(x))M(x)s(x)\| = \lambda_{\max}(\phi(x)).$$

Mais en regardant le membre de droite de l'égalité (27), par l'affirmation (ii) et l'invariance de  $\lambda_{\min}$  par  $\phi$ , nous avons  $\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_{n+1}(x)s(x)\| = \lambda_{\min}(\phi(x)) \neq \lambda_{\max}(\phi(x))$ , ce qui donne une contradiction. L'affirmation (iv) en découle, ainsi que le théorème d'Oseledets 2.12. □

**Exercice E.7.** Soient  $(X, \mathcal{B}, \mu, \phi)$  et  $(X', \mathcal{B}', \mu', \phi')$  deux systèmes dynamiques probabilisés et  $h : X \rightarrow X'$  une semi-conjugaison. Soit  $M' : X' \rightarrow \text{GL}_d(\mathbb{K})$  une application mesurable telle que les applications  $\ln^+ \|(M')^\pm\|$  soient  $\mu'$ -intégrables. Considérons l'application  $M = M' \circ h : X \rightarrow \text{GL}_d(\mathbb{K})$ . Notons  $\lambda'_1(y), \dots, \lambda'_{k'(y)}(y)$  et  $V'_1(y), \dots, V'_{k'(y)}(y)$  les exposants de Lyapounov et espaces de Lyapounov de  $\mu'$ -presque tout  $y \in X'$  pour le produit dynamique de matrice associé à  $(\phi', M')$ . Notons  $\lambda_1(x), \dots, \lambda_{k(x)}(x)$  et  $V_1(x), \dots, V_{k(x)}(x)$  ceux de  $\mu$ -presque tout  $x \in X$  pour  $(\phi, M)$ . Montrer que pour  $\mu$ -presque tout, nous avons  $k = k' \circ h$ ,  $\lambda_i = \lambda'_i \circ h$  et  $V_i = V'_i \circ h$ .

**Exercice E.8.** Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé. Deux applications mesurables  $M, M' : X \rightarrow \text{GL}_d(\mathbb{K})$  sont dites *conjuguées* pour  $\phi$  s'il existe une application mesurable  $A : X \rightarrow \text{GL}_d(\mathbb{K})$  (appelée une *conjugaison* de cocycles linéaires entre  $M$  et  $M'$ ) telle que

- (1) les applications  $\ln \|A\|$  et  $\ln \|A^{-1}\|$  sont intégrables,
- (2) nous avons  $\mu$ -presque partout l'égalité  $M' = A \circ \phi M A^{-1}$ .

Montrer que la relation « être conjuguée à » est une relation d'équivalence sur l'ensemble des applications mesurables de  $X$  dans  $\text{GL}_d(\mathbb{K})$ . Montrer que les exposants de Lyapounov des produits dynamiques de matrices  $(M_n)_{n \in \mathbb{N}}$  et  $(M'_n)_{n \in \mathbb{N}}$  respectivement associés à  $M$  et à  $M'$  sont presque partout égaux.

## 2.5 Exercices

**Exercice E.9.** Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret. Montrer que les assertions suivantes sont équivalentes.

- (1) Le système dynamique  $(X, \mathcal{B}, \mu, \phi)$  est ergodique.
- (2) (**Hypothèse ergodique de Boltzmann**) Pour tous les éléments  $A, B \in \mathcal{B}$  tels que  $\mu(A), \mu(B) > 0$ , il existe  $n \in \mathbb{N}$  tel que  $\mu(\phi^{-n}(A) \cap B) > 0$ .
- (3) Pour tout  $A \in \mathcal{B}$  tel que  $\mu(A) > 0$ , nous avons  $\mu(\bigcup_{n \in \mathbb{N}} \phi^{-n}(A)) = 1$ .
- (4) Pour tout  $A \in \mathcal{B}$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(\phi^{-k}(A) \cap A) = \mu(A)^2 .$$

- (5) (**Propriété de mélange en moyenne de Césaro**)<sup>75</sup> Pour tous les  $A, B \in \mathcal{B}$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(\phi^{-k}(A) \cap B) = \mu(A)\mu(B) .$$

- (6) Il existe des  $\pi$ -systèmes  $\mathcal{E}$  et  $\mathcal{E}'$  engendrant la tribu  $\mathcal{B}$  tels que pour tous les  $A \in \mathcal{E}$  et  $B \in \mathcal{E}'$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(\phi^{-k}(A) \cap B) = \mu(A)\mu(B) .$$

- (7) Pour tout  $f \in \mathbb{L}^1(X, \mu)$  et pour  $\mu$ -presque tout  $x \in X$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ \phi^k(x) = \int_X f d\mu .$$

- (8) Il existe une partie  $E'$  engendrant un sous-espace vectoriel dense de  $\mathbb{L}^1(X, \mu)$  telle que pour tout  $f \in E'$  et pour  $\mu$ -presque tout  $x \in X$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ \phi^k(x) = \int_X f d\mu .$$

---

<sup>75</sup>. Voir la partie 3.1 pour une explication de la terminologie.

- (9) (**Convergence en moyenne de Césaro des corrélations**)<sup>75</sup> Pour toutes les fonctions  $f, g \in \mathbb{L}^2(X, \mu)$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_X f \circ \phi^k g \, d\mu = \lim_{n \rightarrow +\infty} \int_X S_n f g \, d\mu = \int_X f \, d\mu \int_X g \, d\mu .$$

- (10) Il existe des parties  $E'$  et  $E''$  engendrant des sous-espaces vectoriels denses de  $\mathbb{L}^2(X, \mu)$  telles que pour toutes les  $f \in E'$  et  $g \in E''$ , nous avons

$$\lim_{n \rightarrow +\infty} \int_X S_n f g \, d\mu = \int_X f \, d\mu \int_X g \, d\mu .$$

- (11) Pour toute application  $f : X \rightarrow [0, +\infty]$  mesurable qui n'est pas presque partout nulle, la somme  $\sum_{k=0}^{n-1} f \circ \phi^k$  converge presque partout vers  $+\infty$ .

**Exercice E.10.** Soit  $(X, \mathcal{B}, \phi)$  un système dynamique mesurable, et  $\mu$  et  $\nu$  deux mesures de probabilité invariantes. Montrer que si  $\mu$  est absolument continue par rapport<sup>76</sup> à  $\nu$ , et si  $\nu$  est ergodique, alors  $\mu = \nu$ . En déduire que si  $\mu$  et  $\nu$  sont ergodiques, alors elles sont égales ou étrangères.<sup>77</sup>

**Exercice E.11. (Compléments sur le théorème ergodique de Birkhoff 2.2.)** Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique mesuré à temps discret.<sup>78</sup>

- (1) Soit  $f \in \mathbb{L}^2(X, \mu)$ . Montrer que  $\tilde{f} \in \mathbb{L}^2(X, \mu)$  et  $\|\tilde{f}\|_{\mathbb{L}^2} \leq \|f\|_{\mathbb{L}^2}$ , et que si la mesure  $\mu$  est finie, alors la convergence presque partout de  $(S_n f)_{n \in \mathbb{N}}$  vers  $\tilde{f}$  est aussi une convergence  $\mathbb{L}^2$  :

$$\lim_{n \rightarrow \infty} \|S_n f - \tilde{f}\|_{\mathbb{L}^2} = 0 .$$

- (2) Montrer que si  $(X, \mathcal{B}, \mu, \phi)$  est inversible et si la mesure  $\mu$  est finie, alors la moyenne  $\frac{1}{n} \sum_{k=0}^{n-1} f \circ \phi^{-k}(x)$  de  $f$  sur l'orbite inverse de  $x \in X$  converge aussi, pour  $\mu$ -presque tout  $x$ , vers  $\tilde{f}(x)$ .
- (3) Si  $\mu$  est une mesure de probabilité, montrer que  $\phi$  est ergodique si et seulement si pour tout  $f \in \mathbb{L}^1(X, \mu)$  et pour  $\mu$ -presque tout  $x \in X$  nous avons  $\tilde{f}(x) = \int_X f \, d\mu$ .
- (4) Soit  $f \in \mathbb{L}^1(X, \mu)$ . Montrer que si  $\mu$  est une mesure de probabilité et si  $f$  est un *cobord* pour  $(X, \mathcal{B}, \mu, \phi)$ , c'est-à-dire s'il existe une application mesurable  $u : X \rightarrow \mathbb{C}$  telle que  $f = u \circ \phi - u$  presque partout, alors la suite  $(S_n f)_{n \in \mathbb{N}}$  converge presque partout vers 0.

76. Soient  $\mu'$  et  $\nu'$  deux mesures sur un espace mesurable  $(X', \mathcal{B}')$ . Nous dirons que  $\mu'$  est *absolument continue par rapport* à  $\nu'$  si tout ensemble de mesure nulle pour  $\nu'$  l'est aussi pour  $\mu'$ , et la notation usuelle est  $\mu' \ll \nu'$ . Si  $\mu'$  et  $\nu'$  sont  $\sigma$ -finies, le *théorème de Radon-Nykodym* assure alors l'existence d'une fonction mesurable positive  $f : X \rightarrow \mathbb{R}$  telle que  $d\mu' = f \, d\nu'$ , unique modulo égalité  $\mu'$ -presque partout, notée  $\frac{d\mu'}{d\nu'}$  et appelée la *dérivée de Radon-Nykodym* de  $\mu'$  par rapport à  $\nu'$ .

77. Deux mesures  $\mu'$  et  $\nu'$  sur un espace mesurable  $(X', \mathcal{B}')$  sont *étrangères* (on dit « mutually singular » en anglais) s'il existe un ensemble mesurable de mesure pleine pour l'une et de mesure nulle pour l'autre.

78. l'analogie en temps continu est laissé au lecteur.

**Exercice E.12.** Soient  $X$  un espace topologique à base dénombrable d'ouverts<sup>79</sup>,  $\mu$  une mesure de probabilité borélienne sur  $X$  de support<sup>80</sup> égal à  $X$ , et  $\phi : X \rightarrow X$  une application continue qui préserve  $\mu$ . Montrer que si  $\phi$  est ergodique, alors  $\mu$ -presque toute orbite positive du système dynamique topologique à temps discret  $(X, \phi)$  est dense dans  $X$  (c'est-à-dire, pour  $\mu$ -presque tout  $x$  dans  $X$ , l'adhérence de l'orbite positive de  $x$  est l'espace  $X$  tout entier :  $\overline{\{\phi^n(x) : n \in \mathbb{N}\}} = X$ ).

**Exercice E.13. (Transformation induite)** Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret. Soit  $A$  une partie mesurable de mesure non nulle, notons  $\tau_A : X \rightarrow (\mathbb{N} \setminus \{0\}) \cup \{+\infty\}$  son premier temps de retour (voir la partie 1.2). Notons  $\mathcal{B}_A = \{B \cap A : B \in \mathcal{B}\}$  la  $\sigma$ -algèbre induite de  $\mathcal{B}$  sur  $A$ . Rappelons que la *mesure induite*<sup>81</sup> de  $\mu$  sur  $A$  est la mesure de probabilité  $\mu_A$  sur  $(A, \mathcal{B}_A)$  définie par  $B' \mapsto \frac{\mu(B')}{\mu(A)}$ . La *transformation induite* de  $\phi$  dans  $A$  est l'application  $\phi_A : A \rightarrow A$  définie par<sup>82</sup>

$$\phi_A : x \mapsto \begin{cases} \phi^{\tau_A(x)}(x) & \text{si } \tau_A(x) \neq +\infty \\ x & \text{sinon.} \end{cases}$$

- (1) Montrer que la transformation induite  $\phi_A$  est mesurable pour  $\mathcal{B}_A$  et préserve la mesure induite  $\mu_A$ . Le système dynamique  $(A, \mathcal{B}_A, \mu_A, \phi_A)$  est appelé le *système dynamique induit* de  $(X, \mathcal{B}, \mu, \phi)$  sur  $A$ .
- (2) Montrer que si  $(X, \mathcal{B}, \mu, \phi)$  est ergodique alors  $(A, \mathcal{B}_A, \mu_A, \phi_A)$  l'est.<sup>83</sup> Montrer que si  $\mu(X \setminus \bigcup_{n \in \mathbb{N}} \phi^{-n}(A)) = 0$  (ou de manière équivalente si  $\mu(\tau_A^{-1}(+\infty)) = 0$ ) et si  $(A, \mathcal{B}_A, \mu_A, \phi_A)$  est ergodique alors  $(X, \mathcal{B}, \mu, \phi)$  l'est.<sup>84</sup>
- (3) Montrer que si  $(X, \mathcal{B}, \mu, \phi)$  est ergodique, alors pour toute partie mesurable  $A$  de mesure non nulle, la suite des temps de retour  $(\tau_A^k)_{k \in \mathbb{N} \setminus \{0\}}$  dans  $A$  (voir la partie 1.2 pour sa définition) vérifie en presque tout point de  $X$  que

$$\lim_{k \rightarrow +\infty} \frac{\tau_A^k}{k} = \frac{1}{\mu(A)}.$$

- (4) Soient  $\mathbb{K} = \mathbb{R}, \mathbb{C}$ ,  $d \in \mathbb{N} \setminus \{0\}$ ,  $\| \cdot \|$  une norme sur  $\mathcal{M}_d(\mathbb{K})$  et  $M : X \rightarrow \text{GL}_d(\mathbb{K})$  une application mesurable telle que  $\ln^+ \|M\|$  et  $\ln^+ \|M^{-1}\|$  soient intégrables. Notons  $\lambda_1(x), \dots, \lambda_{k(x)}(x)$  et  $V_1(x), \dots, V_{k(x)}(x)$  les exposants de Lyapounov et espaces de Lyapounov de  $\mu$ -presque tout  $x \in X$  pour le produit dynamique de matrice associé à  $(\phi, M)$ . Supposons que  $\tau_A \in \mathbb{L}^1(X, \mu)$  et, quitte à enlever à  $A$  une partie mesurable

79. Un espace topologique  $X$  est à *base dénombrable d'ouverts* s'il existe un ensemble dénombrable  $\mathcal{B}$  d'ouverts de  $X$  tel que tout ouvert de  $X$  soit une union d'éléments de  $\mathcal{B}$ . Par exemple, tout espace topologique métrisable séparable est à base dénombrable d'ouverts. Rappelons qu'un espace topologique est *séparable* s'il admet une partie dénombrable dense. Tout espace topologique métrisable compact est séparable.

80. Le *support* d'une mesure positive borélienne est le plus petit fermé en dehors duquel la mesure est nulle ; c'est l'ensemble des points dont tout voisinage ouvert est de mesure non nulle.

81. C'est la restriction à  $\mathcal{B}_A$  de la mesure conditionnelle  $B \mapsto \mu(B | A) = \frac{\mu(B \cap A)}{\mu(A)}$  sur  $\mathcal{B}$  sachant  $A$ .

82. Par le théorème de récurrence de Poincaré 1.1, l'ensemble des points  $x$  de  $A$  tels que  $\tau_A(x) = +\infty$  est mesurable de mesure nulle, donc la manière de définir  $\phi_A$  sur cet ensemble est matière de convention, elle n'a aucune importance du point de vue de la théorie ergodique.

83. Prendre garde que ce résultat devient incorrect pour des propriétés plus fortes que l'ergodicité.

84. Par contre, la propriété de mélange n'est pas toujours préservée par passage à un système dynamique induit, voir l'exercice E.41.



de mesure nulle, que  $\tau_A(X) \subset \mathbb{N} \setminus \{0\}$ . Notons  $M' : A \rightarrow \text{GL}_d(\mathbb{K})$  l'application  $x \mapsto M_{\tau_A(x)}(x)$ .

- a) Montrer que les applications  $\ln^+ \|M'\|$  et  $\ln^+ \|M'^{-1}\|$  sont  $\mu_A$ -intégrables.
- b) Montrer que pour  $\mu_A$ -presque tout  $x \in A$ , la limite  $a(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \tau_A \circ \phi^k(x)$  existe dans  $[1, +\infty[$ .
- c) Pour  $\mu_A$ -presque tout  $x \in A$ , montrer que les exposants de Lyapounov en  $x$  pour le produit dynamique de matrice associé à  $(\phi_A, M')$  sont  $\lambda'_1 = a(x)\lambda_1(x), \dots, \lambda'_{k'(x)}(x) = a(x)\lambda_{k(x)}(x)$  et que ses espaces de Lyapounov sont  $V'_1 = V_1(x), \dots, V'_{k'(x)}(x) = V_{k(x)}(x)$ .

**Exercice E.14. (Transformation de Kakutani)** Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret inversible. Soit  $\tau : X \rightarrow (\mathbb{N} \setminus \{0\}) \cup \{\infty\}$  une application mesurable, dont l'intégrale vérifie  $\int_X \tau d\mu = 1$ , appelée une *fonction de hauteur* (normalisée) sur  $X$ . Soit

$$X_\tau = \{(x, n) \in X \times \mathbb{N} : n < \tau(x)\}$$

muni de la  $\sigma$ -algèbre  $\mathcal{B}_\tau$  induite de la  $\sigma$ -algèbre produit de  $X \times \mathbb{N}$ . Soit  $\phi_\tau : X_\tau \rightarrow X_\tau$  l'application définie par<sup>85</sup>

$$(x, n) \mapsto \begin{cases} (x, n+1) & \text{si } n+1 < \tau(x) \\ (\phi(x), 0) & \text{sinon,} \end{cases}$$

appelée la *transformation de Kakutani* sur  $X$  de fonction de hauteur  $\tau$ .

- (1) Montrer que la transformation de Kakutani  $\phi_\tau$  est mesurable pour  $\mathcal{B}_\tau$ .
- (2) Montrer qu'il existe sur l'espace mesurable  $(X_\tau, \mathcal{B}_\tau)$  une unique mesure de probabilité  $\mu_\tau$ , invariante par  $\phi_\tau$  et telle que la mesure induite (au sens de l'exercice E.13) de  $\mu_\tau$  sur  $X \times \{0\}$  soit la mesure  $\mu$ . L'espace mesuré  $(X_\tau, \mathcal{B}_\tau, \mu_\tau)$  est appelé la *tour de Kakutani* sur  $(X, \mathcal{B}, \mu)$  de fonction de hauteur  $\tau$ . Faire un dessin. Le système dynamique mesuré  $(X_\tau, \mathcal{B}_\tau, \mu_\tau, \phi_\tau)$  est appelé la *suspension* du système dynamique à temps discret  $(X, \mathcal{B}, \mu, \phi)$  de fonction de hauteur  $\tau$ .
- (3) Montrer que si  $(X, \mathcal{B}, \mu, \phi)$  est ergodique, alors  $(X_\tau, \mathcal{B}_\tau, \mu_\tau, \phi_\tau)$  l'est aussi.
- (4) Soient  $(Y, \mathcal{C}, \nu, \psi)$  un système dynamique probabilisé à temps discret inversible et ergodique, et  $A$  une partie mesurable de mesure non nulle de  $Y$  telle que  $\mu(\tau_A^{-1}(+\infty)) = 0$ . Montrer que le système dynamique mesuré  $(Y, \mathcal{C}, \nu, \psi)$  est conjugué au système dynamique mesuré suspension du système dynamique induit  $(A, \mathcal{C}_A, \nu_A, \psi_A)$  de  $(Y, \mathcal{C}, \nu, \psi)$  sur  $A$  pour la fonction de hauteur égale au premier temps de retour  $\tau_A$ .

**Exercice E.15. (Produits croisés)** Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique mesuré à temps discret. Soient  $G$  un groupe topologique localement compact,  $\mathcal{B}_G$  sa  $\sigma$ -algèbre borélienne et  $\mu_G$  une mesure de Haar de  $G$ . Soit  $f : X \rightarrow G$  une application mesurable. Considérons l'application  $\phi_f : X \times G \rightarrow X \times G$  définie par

$$(x, g) \mapsto (\phi(x), f(x)g).$$

- (1) Montrer que  $\phi_f$  est mesurable pour la  $\sigma$ -algèbre produit  $\mathcal{B} \otimes \mathcal{B}_G$ , et préserve la mesure produit  $\mu \otimes \mu_G$ . Le quadruplet  $(X \times G, \mathcal{B} \otimes \mathcal{B}_G, \mu \otimes \mu_G, \phi_f)$  est le *système dynamique produit croisé* de  $(X, \mathcal{B}, \mu, \phi)$  par  $f$ .

<sup>85</sup> La normalisation  $\int_X \tau d\mu = 1$  de la fonction de hauteur  $h$  implique que l'ensemble mesurable  $\tau^{-1}(+\infty) = \{x \in X : \tau(x) = +\infty\}$  est de mesure nulle pour  $\mu$ . Donc la définition de la transformation de Kakutani en un point  $(x, n)$  où  $\tau(x) = +\infty$  est matière de convention, ici nous prenons  $(x, n) \mapsto (x, n+1)$ .

- (2) Montrer que le système dynamique mesurable  $(X, \mathcal{B}, \phi)$  est un facteur du système dynamique mesurable  $(X \times G, \mathcal{B} \otimes \mathcal{B}_G, \phi_f)$ , et que si  $G$  est compact (et  $\mu_G$  normalisée), alors le système dynamique mesuré  $(X, \mathcal{B}, \mu, \phi)$  est un facteur du système dynamique mesuré  $(X \times G, \mathcal{B} \otimes \mathcal{B}_G, \mu \otimes \mu_G, \phi_f)$ . Si  $G$  est abélien, montrer que pour tous les  $n \in \mathbb{N}$  et  $(x, g) \in X \times G$ , nous avons

$$(\phi_f)^n(x, g) = (\phi^n(x), g + f(x) + f \circ \phi(x) + \dots + f \circ \phi^{n-1}(x)). \quad (28)$$

- (3) (**Théorème d'Anzai**) Si  $\phi$  est ergodique et si  $G$  est le cercle  $\mathbb{R}/\mathbb{Z}$ , montrer, par exemple en utilisant le développement de Fourier en la seconde variable des fonctions réelles définies sur  $X \times G$  (voir la partie 4 pour plus d'information), que le système dynamique produit croisé est ergodique si et seulement s'il existe  $p \in \mathbb{Z} \setminus \{0\}$  tel que la fonction  $pf : X \rightarrow G$  soit un *cobord* à valeurs dans  $G$ , c'est-à-dire qu'il existe une application mesurable  $\theta : X \rightarrow G$  telle que presque partout

$$pf = \theta \circ \phi - \theta.$$

**Exercice E.16.** Soit  $x \in \{1, \dots, 9\}$  un chiffre non nul. Quelle est la fréquence d'apparition de  $x$  comme premier chiffre dans l'écriture décimale des puissances  $2^n$  de 2 pour  $n \in \mathbb{N}$  ?

**Exercice E.17.** Un système dynamique probabilisé à temps discret  $(X, \mathcal{B}, \mu, \phi)$  est dit *totalelement ergodique* si pour tout  $n \in \mathbb{N} \setminus \{0\}$ , le système dynamique probabilisé à temps discret  $(X, \mathcal{B}, \mu, \phi^n)$  est ergodique.

- (1) Donner un exemple de système dynamique probabilisé à temps discret ergodique non totalement ergodique.
- (2) Montrer que le système dynamique  $(X, \mathcal{B}, \mu, \phi)$  est totalement ergodique si et seulement s'il n'admet pas de facteur fini non trivial (c'est-à-dire dont le support de la mesure n'est pas un singleton).

**Exercice E.18. (Minoration d'exposant de Lyapounov)** Notons  $\|\cdot\|_\infty$  la norme  $\|(a_{ij})_{1 \leq i, j \leq 2}\|_\infty = \max_{1 \leq i, j \leq 2} |a_{ij}|$  sur l'espace vectoriel réel  $\mathcal{M}_2(\mathbb{R})$  des matrices  $2 \times 2$  réelles. Soient  $c > 0$  et  $\alpha \in \mathbb{R} - \mathbb{Q}$  un irrationnel. Notons  $X = \mathbb{R}/\mathbb{Z}$  le cercle muni de sa mesure de Haar normalisée, et  $\phi : X \rightarrow X$  la translation irrationnelle définie par  $x \mapsto x + \alpha$ . Pour tout  $x \in X$ , notons  $R_x = \begin{pmatrix} \cos(2\pi x) & -\sin(2\pi x) \\ \sin(2\pi x) & \cos(2\pi x) \end{pmatrix}$  la rotation d'angle  $2\pi x$  et  $H = \begin{pmatrix} c & 0 \\ 0 & c^{-1} \end{pmatrix}$ . Considérons l'application continue  $M : X \rightarrow \text{SL}_2(\mathbb{R})$  définie par  $x \mapsto HR_x$  et  $(M_n)_{n \in \mathbb{N}}$  le produit dynamique de matrices à gauche associé, de sorte que pour tous les  $n \in \mathbb{N}$  et  $x \in X$ ,

$$M_n(x) = M(\phi^{n-1}(x)) M(\phi^{n-2}(x)) \dots M(\phi(x)) M(x).$$

Notons  $\lambda_{\max}$  son plus grand exposant de Lyapounov.

- (1) Montrer que presque partout

$$\lambda_{\max} = \lim_{n \rightarrow +\infty} \int_0^1 \frac{1}{n} \ln \|HR_{x+(n-1)\alpha} \dots HR_{x+\alpha} HR_x\|_\infty dx.$$

- (2) Pour tout  $z \in \mathbb{C}$ , posons  $\tilde{R}(z) = \begin{pmatrix} (z^2+1)/2 & (-z^2+1)/(2i) \\ (z^2-1)/(2i) & (z^2+1)/2 \end{pmatrix}$ . Exprimer  $R_x$  en fonction de  $\tilde{R}(e^{2i\pi x})$  pour tout  $x \in X$ .

- (3) Montrer qu'il existe une suite  $(f_n)_{n \in \mathbb{N}}$  de fonctions sous-harmoniques<sup>86</sup> de  $\mathbb{C}$  dans  $\mathbb{R}$  telles que  $\lambda_{\max} = \lim_{n \rightarrow +\infty} \frac{1}{n} \int_0^1 f_n(e^{2i\pi x}) dx$  presque partout.
- (4) En déduire que presque partout  $\lambda_{\max} \geq \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|(H\tilde{R}(0))^n\|_{\infty}$ , puis que presque partout  $\lambda_{\max} \geq \ln \frac{c+c^{-1}}{2}$ .

## 2.6 Indications pour la résolution des exercices

**Correction de l'exercice E.7.** Il est immédiat que  $M$  est mesurable. Puisque  $h_*\mu = \mu'$ , nous avons  $\int_X \ln^+ \|M^{\pm}\| d\mu = \int_{X'} \ln^+ \|(M')^{\pm}\| d\mu' < +\infty$ , donc  $k, \lambda_1, \dots, \lambda_k$  et  $V_1, \dots, V_k$  sont bien définis  $\mu$ -presque partout par le théorème d'Oseledets 2.12. Pour tout  $n \in \mathbb{N}$ , puisque  $h \circ \phi = \phi' \circ h$ , nous avons  $M_n = M'_n \circ h$ , donc pour tout  $v \in \mathbb{K}^d \setminus \{0\}$ ,

$$\frac{1}{n} \ln \|M_n v\| = \frac{1}{n} \ln \|(M'_n \circ h)v\| .$$

Le résultat en découle.

**Correction de l'exercice E.8.** Il est immédiat que l'application constante  $x \mapsto I_d$  est une conjugaison entre  $M$  et  $M$ , que si  $A$  est une conjugaison entre  $M$  et  $M'$ , alors  $A^{-1}$  est une conjugaison entre  $M'$  et  $M$ , et que si  $B$  est une conjugaison entre  $M'$  et  $M''$ , alors  $BA$  est une conjugaison entre  $M$  et  $M''$ .

Quitte à enlever à  $X$  un ensemble de mesure nulle, notons  $\lambda_1, \dots, \lambda_k : X \rightarrow \mathbb{R}$  les exposants de Lyapounov (et  $k : X \rightarrow \llbracket 1, d \rrbracket$  leur nombre) et  $V_1, \dots, V_k : X \rightarrow \mathbb{R}$  les sous-espaces de Lyapounov pour  $M$ .

Pour tous les  $n \in \mathbb{N}$  et  $x \in X$ , par la propriété (2) d'une conjugaison de cocycles linéaires et par simplifications, nous avons

$$\begin{aligned} M'_n(x) &= M'(\phi^{n-1}(x)) M'(\phi^{n-2}(x)) \dots M'(\phi(x)) M'(x) \\ &= A(\phi^n(x)) M(\phi^{n-1}(x)) M(\phi^{n-2}(x)) \dots M(\phi(x)) M(x) A(x)^{-1} \\ &= A(\phi^n(x)) M_n(x) A(x)^{-1} . \end{aligned}$$

Posons  $V'_1(x) = A(x)V_1(x), \dots, V'_k(x) = A(x)V_k(x)$ . Nous avons  $v' \in V'_i(x) \setminus V'_{i+1}(x)$  si et seulement s'il existe  $v \in V_i(x) \setminus V_{i+1}(x)$  tel que  $v' = A(x)v$ , et alors  $M'_n(x)v' = A(\phi^n(x))M_n(x)v$ . Par les propriétés de la norme d'opérateur, nous avons alors

$$\left| \frac{1}{n} \ln \|M'_n(x)v'\| - \frac{1}{n} \ln \|M_n(x)v\| \right| \leq \max \left\{ \frac{1}{n} \ln \|A(\phi^n(x))\|, \frac{1}{n} \ln \|A^{-1}(\phi^n(x))\| \right\} .$$

Par le lemme 2.9 (ou 2.13 en fonction des préférences pour la démonstration) et la propriété (1) d'une conjugaison de cocycles linéaires, nous avons presque partout

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|A(\phi^n(x))\| = 0$$

---

86. Rappelons qu'une fonction continue  $f : \mathbb{C} \rightarrow \mathbb{R}$  est *sous-harmonique* si la valeur de  $f$  en tout point  $z_0 \in \mathbb{C}$  est inférieure à la moyenne de  $f$  sur tout cercle centré en  $z_0$  :

$$\forall z_0 \in \mathbb{C}, \forall r > 0, \quad f(z_0) \leq \int_0^1 f(z_0 + r e^{2i\pi t}) dt .$$

Rappelons que si  $u : \mathbb{C} \rightarrow \mathbb{C}$  est holomorphe, alors  $z \mapsto \ln |u(z)|$  est sous-harmonique, et qu'une fonction  $f : \mathbb{C} \rightarrow \mathbb{R}$  qui est le maximum d'un nombre fini de fonctions sous-harmoniques est sous-harmonique.

et de même pour  $A^{-1}$ .

Ceci montre que  $M$  et  $M'$  ont presque partout les mêmes exposants de Lyapounov, ainsi que le fait que presque partout les sous-espaces de Lyapounov pour  $M'$  sont les images par  $A$  des sous-espaces de Lyapounov pour  $M$ .

**Correction de l'exercice E.9.** Les implications (5)  $\Rightarrow$  (4), (5)  $\Rightarrow$  (6), (7)  $\Rightarrow$  (8) sont immédiates.

(2)  $\Rightarrow$  (1). Si  $A \in \mathcal{B}$  est tel que  $\phi^{-1}(A) = A$  et  $\mu(A) > 0$ , alors pour tout  $n \in \mathbb{N}$ , nous avons  $\mu(\phi^{-n}(A) \cap {}^c A) = \mu(A \cap {}^c A) = 0$ , donc  $\mu({}^c A) = 0$  par l'assertion (2).

(1)  $\Rightarrow$  (3). Pour tout  $A \in \mathcal{B}$  tel que  $\mu(A) > 0$ , si  $B = \bigcup_{n \in \mathbb{N}} \phi^{-n}(A)$ , alors  $\mu(B) \geq \mu(A) > 0$  et  $\phi^{-1}(B) = B$  donc  $\mu(B) = 1$  par ergodicité.

(3)  $\Rightarrow$  (2). Pour tous les  $A, B \in \mathcal{B}$  tels que  $\mu(A), \mu(B) > 0$ , l'assertion (3) dit que  $\mu(\bigcup_{n \in \mathbb{N}} \phi^{-n}(A)) = 1$ . Donc  $\mu(B \cap \bigcup_{n \in \mathbb{N}} \phi^{-n}(A)) = \mu(B) > 0$  et comme une union dénombrable d'ensemble de mesure nulle est de mesure nulle, il existe  $n \in \mathbb{N}$  tel que  $\mu(B \cap \phi^{-n}(A)) > 0$ .

(3)  $\Rightarrow$  (1). (Argument direct) Si  $A \in \mathcal{B}$  est tel que  $\phi^{-1}(A) = A$  et  $\mu(A) > 0$ , alors l'assertion (3) implique que  $\mu(A) = \mu(\bigcup_{n \in \mathbb{N}} \phi^{-n}(A)) = 1$ .

(4)  $\Rightarrow$  (1). Si  $A$  est une partie mesurable invariante, alors  $\frac{1}{n} \sum_{k=0}^{n-1} \mu(\phi^{-k}(A) \cap A) = \mu(A)$ . Donc l'assertion (4) implique que  $\mu(A) = \mu(A)^2$ , c'est-à-dire que  $\mu(A)$  vaut 0 ou 1.

(1)  $\Rightarrow$  (5). La fonction  $f = \mathbb{1}_A$  est intégrable. La convergence  $\mathbb{L}^1$  de  $S_n f$  vers la fonction constante  $\int_X f d\mu$  (voir le théorème de Birkhoff 2.2 (4) et (6)) montre que

$$\frac{1}{n} \sum_{k=0}^{n-1} \mu(\phi^{-k}(A) \cap B) = \int_B S_n f d\mu \xrightarrow{n \rightarrow +\infty} \int_B \int_X f d\mu d\mu = \mu(A) \mu(B).$$

(1)  $\Rightarrow$  (7). C'est l'assertion (6) du théorème de Birkhoff 2.2.

(7)  $\Rightarrow$  (1). Si  $A \in \mathcal{B}$  est tel que  $\phi^{-1}(A) = A$  et  $\mu(A) > 0$ , soit  $f = \mathbb{1}_A$ , qui est mesurable, intégrable et invariante. Par l'assertion (7), pour  $\mu$ -presque tout  $x \in X$ , nous avons  $f(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ \phi^k(x) = \int_X f d\mu$ . Donc  $f$  est presque partout constante, et égale à 1 sur l'ensemble de mesure non nulle  $A$ , donc est presque partout égale à 1. D'où  $\mu(A) = 1$ .

**Correction de l'exercice E.10.** Si  $\mu$  et  $\nu$  sont deux mesures de probabilité invariantes, si  $\mu$  est absolument continue par rapport à  $\nu$ , alors la dérivée de Radon-Nikodym  $\frac{d\mu'}{d\nu'}$  est mesurable et  $\nu$ -presque partout invariante. Donc si  $\nu$  est ergodique, elle est  $\nu$ -presque partout constante, égale à 1 car  $\mu$  et  $\nu$  sont deux mesures de probabilité.

Si  $\mu$  et  $\nu$  sont deux mesures de probabilité invariantes ergodiques non égales, alors il existe une partie mesurable  $A$  telle que  $\mu(A) \neq \nu(A)$ . La fonction  $f = \mathbb{1}_A$  est intégrable. Par le théorème de Birkhoff 2.2 (6), il existe des ensembles  $X_\mu$  et  $X_\nu$  mesurables de mesures plines pour  $\mu$  et  $\nu$  respectivement tels que  $\lim_{n \rightarrow +\infty} S_n f(x) = \mu(A)$  pour tout  $x \in X_\mu$  et  $\lim_{n \rightarrow +\infty} S_n f(x) = \nu(A)$  pour tout  $x \in X_\nu$ . Comme  $\mu(A) \neq \nu(A)$ , ceci montre que les parties  $X_\mu$  et  $X_\nu$  sont disjointes, donc que  $\mu$  et  $\nu$  sont étrangères.

**Correction de l'exercice E.11.** (1) La démonstration est similaire à celle dans le cas  $\mathbb{L}^1$ .

(2) Si  $\phi$  est inversible, la tribu des invariants  $\mathcal{I}$  de  $\phi$  est égale à celle de  $\phi^{-1}$ , et le résultat découle du fait que  $\tilde{f} = \mathbb{E}(f|\mathcal{I})$ .

(3) Le sens direct est l'assertion (6) du théorème ergodique de Birkhoff 2.2. Réciproquement, soit  $f \in \mathbb{L}^1(X, \mu)$  une application presque partout invariante. Alors pour tout  $n \in \mathbb{N}$  et presque partout, nous avons  $\tilde{f} = S_n f$ , donc par passage à la limite presque partout, nous avons  $f = \tilde{f}$ . L'hypothèse  $f = \int_X f d\mu$  presque partout implique donc que  $f$  est presque partout constante. Par conséquent  $\phi$  est ergodique par la proposition 2.1 (3).

**Correction de l'exercice E.12.** Soit  $\mathcal{C}$  une base dénombrable d'ouverts non vides dans  $X$ . Pour tout élément  $B$  de  $\mathcal{C}$ , la partie  $A_B = \bigcap_{p \in \mathbb{N}} \bigcup_{n \geq p} \phi^{-n}(B)$  (formée des points de  $X$  passant une infinité de fois dans  $B$ ) vérifie  $\phi^{-1}(A_B) = A_B$ . Cette partie est une intersection décroissante de parties de mesure supérieure ou égale à  $\mu(B)$ . Donc  $\mu(A_B) \geq \mu(B) > 0$  puisque le support de  $\mu$  est total. Par ergodicité de  $\phi$ , nous avons donc  $\mu(A_B) = 1$ . L'intersection  $A = \bigcap_{B \in \mathcal{C}} A_B$  est aussi de mesure totale, et les orbites des points de  $A$  sont denses, car elles passent une infinité de fois dans chaque élément de  $\mathcal{C}$ , donc dans chaque ouvert non vide de  $X$ .

**Correction de l'exercice E.13.** (3) L'application  $\phi_A : A \rightarrow A$  de premier retour dans  $A$  de  $\phi$  est ergodique pour  $\mu_A$  par l'assertion (2). Par la formule (9), nous avons par récurrence

$$\forall x \in A, \quad \tau_A^k(x) = \sum_{i=0}^{k-1} \tau_A^1(\phi_A^i(x)),$$

Donc  $\frac{\tau_A^k}{k} = S_k \tau_A^1$  est la  $k$ -ème moyenne de Birkhoff de  $\tau_A^1$  pour la transformation  $\phi_A$ . Le temps de premier retour  $\tau_A = \tau_A^1 : A \rightarrow \mathbb{R}$  est intégrable pour  $\mu_A$ , et le théorème de Kac 1.2, avec l'exercice E.6 qui dit que  $\mu(\tau_A^{-1}(\mathbb{N} \setminus \{0\})) = 1$ , montre, outre cette intégrabilité, que  $\int_A \tau_A d\mu_A = \frac{1}{\mu(A)}$ . Par le théorème de Birkhoff 2.2 (6) appliqué au système dynamique probabilisé induit  $(A, \mathcal{B}_A, \mu_A, \phi_A)$  et à la fonction  $\tau_A : A \rightarrow \mathbb{R}$ , pour presque tout (pour  $\mu_A$  donc pour  $\mu$ ) élément de  $A$ , la limite  $\lim_{k \rightarrow +\infty} \frac{\tau_A^k}{k}$  existe, et vaut  $\frac{1}{\mu(A)}$ .

Comme  $\mu$ -presque tout point  $x$  de  $X \setminus A$  passe par  $A$  en temps  $\tau_A(x)$  fini, comme  $\frac{\tau_A^k(x)}{k} = \frac{\tau_A(x)}{k} + \frac{\tau_A^{k-1}(\phi_A(x))}{k-1} \frac{k-1}{k}$ , le résultat est aussi valable pour presque tout  $x \in X \setminus A$ .

(4) Nous pouvons supposer que  $\| \cdot \|$  est une norme d'opérateur. Remarquons que l'application  $M'$  est mesurable, car  $\tau_A$  et  $M$  le sont.

a) Pour tout  $i \in \mathbb{N} \setminus \{0\}$ , notons  $A_i = \{x \in A : \tau_A(x) = i\}$ , et remarquons que  $A = \bigsqcup_{i \geq 1} A_i$ , que  $\phi_A|_{A_i} = \phi^i|_{A_i}$  et que  $M'|_{A_i} = M|_{A_i}$ . Donc, puisque  $\| \cdot \|$  est une norme d'opérateur par invariance de la mesure  $\mu$  par  $\phi$ , et puisque les parties  $\phi^{-1}(A_i)$  de  $X$  pour  $i \in \mathbb{N} \setminus \{0\}$  et  $k \in \llbracket 0, i-1 \rrbracket$  sont deux à deux disjointes, nous avons

$$\begin{aligned} \int_A \ln^+ \|M'\| d\mu_A &= \sum_{i \geq 1} \frac{1}{\mu(A)} \int_{A_i} \ln^+ \|M_i\| d\mu \leq \frac{1}{\mu(A)} \sum_{i \geq 1} \sum_{k=1}^{i-1} \int_{A_i} \ln^+ \|M \circ \phi^k\| d\mu \\ &= \frac{1}{\mu(A)} \sum_{i \geq 1} \sum_{k=1}^{i-1} \int_{\phi^{-1}(A_i)} \ln^+ \|M\| d\mu \leq \frac{1}{\mu(A)} \int_X \ln^+ \|M\| d\mu. \end{aligned}$$

Donc l'intégrabilité de  $\ln^+ \|M\|$  pour  $\mu$  implique celle de  $\ln^+ \|M'\|$  pour  $\mu_A$ . Le résultat pour  $\ln^+ \|M'^{-1}\|$  est analogue.

b) Ceci découle du théorème de Birkhoff appliqué à l'application  $\tau_A$  qui a été supposée intégrable.

c) Pour tout  $n \in \mathbb{N} \setminus \{0\}$ , par la définition de  $M'$  et par la formule de cocycle multiplicatif (19), nous avons

$$\begin{aligned} M'_n &= M' \circ \phi_A^{n-1} M' \circ \phi_A^{n-2} \dots M' \circ \phi_A M' = M_{\tau_A \circ \phi_A^{n-1}} M_{\tau_A \circ \phi_A^{n-2}} \dots M_{\tau_A \circ \phi_A} M_{\tau_A} \\ &= M_{\sum_{k=0}^{n-1} \tau_A \circ \phi_A^k}. \end{aligned}$$

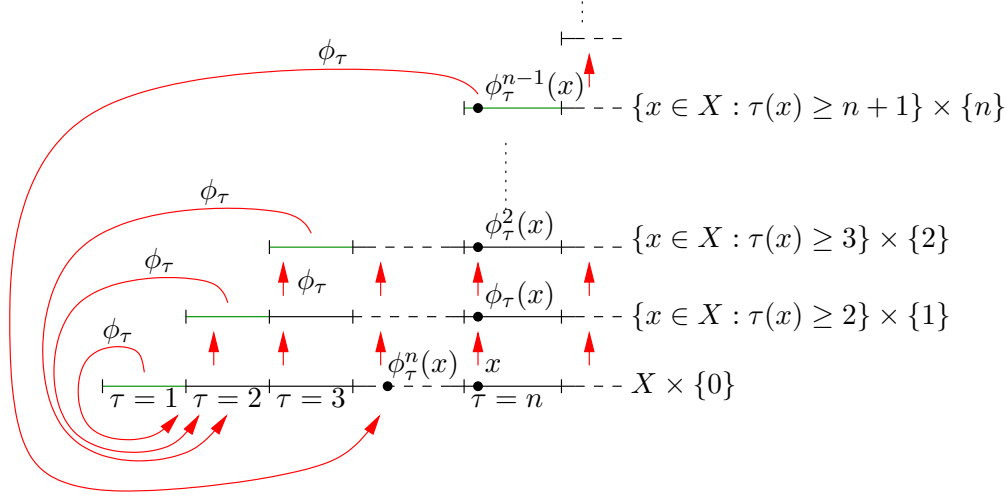
Pour tout  $x \in A$ , pour tout  $v \in V_i(x) \setminus V_{i+1}(x)$ , nous avons donc

$$\frac{1}{n} \ln \|M'_n v\| = \left( \frac{1}{n} \sum_{k=0}^{n-1} \tau_A \circ \phi^k(x) \right) \left( \frac{1}{\sum_{k=0}^{n-1} \tau_A \circ \phi^k(x)} \ln \|M_{\sum_{k=0}^{n-1} \tau_A \circ \phi^k} v\| \right).$$

Le premier facteur du terme de droite converge presque partout vers  $a(x)$  quand  $n \rightarrow +\infty$  par la question b). Le second facteur converge presque partout vers  $\lambda_i(x)$  par le théorème d'Oseledets 2.12. Le résultat en découle.

**Correction de l'exercice E.14.**

(2)



**Correction de l'exercice E.15.** (1) L'application  $\phi_f$  est mesurable car chacune de ses composantes l'est. Pour montrer que la mesure produit  $\mu \otimes \mu_G$  est  $\phi_f$ -invariante, il suffit (voir la formule 5) de vérifier son invariance sur les fonctions à variables séparables  $h : (x, g) \mapsto u(x)v(g)$  avec  $u \in \mathbb{L}^1(X, \mu)$  et  $v \in \mathbb{L}^1(G, \mu \otimes \mu_G)$ , qui engendrent un sous-espace vectoriel dense de  $\mathbb{L}^1(X \times G, \mu \otimes \mu_G)$ . Or puisque la mesure de Haar  $\mu_G$  de  $G$  est invariante par translation par  $f(x)$  pour tout  $x \in X$  et puisque  $\mu$  est  $\phi$ -invariante, Rappelons que pour tout  $h \in G$ , l'application  $\tau_h$  est la translation à gauche par  $h$  dans  $G$ , nous avons par le théorème de Fubini

$$\begin{aligned} \int_{X \times G} h \circ \phi_f d(\mu \otimes \mu_G) &= \int_{X \times G} u \circ \phi(x) v \circ \tau_{f(x)}(g) d(\mu \otimes \mu_G)(x, g) \\ &= \int_X u \circ \phi(x) \left( \int_G v \circ \tau_{f(x)}(g) d\mu_G(g) \right) d\mu(x) = \int_X u \circ \phi(x) \left( \int_G v d\mu_G \right) d\mu(x) \\ &= \left( \int_X u d\mu \right) \left( \int_G v d\mu_G \right) = \int_{X \times G} h d(\mu \otimes \mu_G). \end{aligned}$$

(2) L'application  $\text{pr}_1 : (x, g) \mapsto x$  de  $X \times G$  dans  $X$  est mesurable surjective, et vérifie  $\text{pr}_1 \circ \phi_f = \phi \circ \text{pr}_1$  par la définition de  $\phi_f$ . Ceci montre la propriété de facteur mesurable. Si  $\mu_G$  est une mesure de probabilité, alors  $(\text{pr}_1)_*(\mu \otimes \mu_G) = \mu$ . Ceci montre la propriété de facteur mesuré. La formule (28) est immédiate par récurrence.

(3) Nous raisonnons par contraposition. Supposons que  $\phi_f$  ne soit pas ergodique. Soit  $h \in \mathbb{L}^2(X \times G)$  une fonction presque partout  $\phi_f$ -invariante et non presque partout constante. Par le théorème de Fubini, l'application  $t \mapsto h(x, t)$  appartient à  $\mathbb{L}^2(\mathbb{R}/\mathbb{Z})$  pour presque tout  $x \in X$ . Pour tout  $p \in \mathbb{Z}$ , notons  $c_p : X \rightarrow \mathbb{C}$  le  $p$ -ème coefficient de Fourier en la seconde variable de la fonction  $h$ , défini par

$$c_p : x \mapsto \int_0^1 h(x, t) e^{-2\pi i p t} dt .$$

Par l'invariance presque partout par  $\phi_f$  de  $h$ , par l'invariance de la mesure de Haar de  $\mathbb{R}/\mathbb{Z}$  et par l'unicité des coefficients de Fourier des fonctions  $\mathbb{L}^2$  sur le cercle, pour presque tout  $x \in X$ , nous avons

$$c_p(\phi(x)) e^{2\pi i p f(x)} = c_p(x) . \quad (29)$$

Pour  $p = 0$ , cette égalité dit que l'application  $c_0$  est presque partout invariante par  $\phi$ . Donc par l'ergodicité de  $\phi$ , elle est presque partout constante. Puisque  $h$  n'est pas presque partout constante, il existe  $p \in \mathbb{Z} - \{0\}$  tel que la fonction  $c_p$  n'est pas presque partout nulle. La formule (29) implique que  $|c_p(\phi(x))| = |c_p(x)|$ . De nouveau par ergodicité de  $\phi$ , il existe donc  $C_p > 0$  tel que l'application  $x \mapsto |c_p(x)|$  soit presque partout égale à  $C_p$ . Par le théorème du relèvement, il existe donc une application  $\theta : X \rightarrow \mathbb{R}/\mathbb{Z}$  mesurable telle que pour presque tout  $x \in X$ , nous ayons  $c_p(x) = C_p e^{2\pi i \theta(x)}$ . La formule (29) montre alors que pour presque tout  $x \in X$  et modulo 1, nous avons  $\theta(\phi(x)) + p f(x) = \theta(x)$ .

Réciproquement, s'il existe une application mesurable  $\theta : X \rightarrow G$  et  $p \in \mathbb{Z} \setminus \{0\}$  tels que  $p f = \theta \circ \phi - \theta$  presque partout, alors l'application  $h : (x, t) \mapsto e^{2\pi i(\theta(x) + p t)}$  est presque partout  $\phi_f$ -invariante et non presque partout constante, et donc  $\phi_f$  n'est pas ergodique.

**Correction de l'exercice E.16.** Soit  $k \in \llbracket 1, 9 \llbracket$  un chiffre non nul. Pour tout  $n \in \mathbb{N}$ , notons  $a_n$  le premier chiffre de l'écriture décimale de  $2^n$ . Montrons que

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \text{Card}\{n \in \llbracket 0, N \llbracket : r_n = k\} = \log_{10} \frac{k+1}{k} .$$

Nous avons

$$\begin{aligned} a_n = k &\iff \exists \ell \in \mathbb{N}, \quad k 10^\ell \leq 2^n < (k+1)10^\ell \\ &\iff n \log_{10}(2) \in [\log_{10}(k), \log_{10}(k+1)[ + \mathbb{N} \end{aligned}$$

Le réel  $\alpha = \log_{10}(2) = \frac{\ln 2}{\ln 10}$  est irrationnel. Par l'unique ergodicité des rotations irrationnelles (voir l'exemple à la fin de la partie 4.2 et le critère de Weyl 4.4 pour le passage des fonctions continues aux fonctions caractéristique d'intervalle), la fréquence des passages dans l'intervalle  $[\log_{10}(k), \log_{10}(k+1)[$  des images de 0 par la rotation irrationnelle d'angle  $\alpha$  converge vers la longueur de cet intervalle.

**Correction de l'exercice E.17.** (1) Soient  $k \in \mathbb{N} \setminus \{0, 1\}$ ,  $X$  l'ensemble fini  $\mathbb{Z}/k\mathbb{Z}$  muni de sa tribu discrète  $\mathcal{B} = \mathcal{P}(X)$ ,  $\mu$  l'équiprobabilité sur  $X$  et  $\phi : x \rightarrow x + 1 \pmod k$  le décalage de 1. Alors ce système dynamique est ergodique mais pas totalement ergodique :

pour tout  $n \in \mathbb{N} \setminus \{0\}$ , la transformation  $\phi^n$  est ergodique si et seulement si  $n$  est premier avec  $k$ .

(2) Soit  $(X', \mathcal{B}', \mu', \phi')$  un système dynamique probabilisé à temps discret facteur de  $(X, \mathcal{B}, \mu, \phi)$  avec  $X'$  fini et  $h : X \rightarrow X'$  une semi-conjugaison. Puisque  $X'$  est fini, il existe  $k \in \mathbb{N} \setminus \{0\}$  tel que  $(\phi')^k = \text{id}_{X'}$ . Soit  $x' \in X'$  tel que  $0 < \mu'(\{x'\}) < 1$  (qui existe car  $X'$  est fini et la mesure n'est pas de support un singleton). Alors l'ensemble  $h^{-1}(x')$  est un ensemble mesurable  $\phi^k$ -invariant par la propriété de semi-conjugaison, qui n'est pas de mesure nulle ou pleine car  $h$  préserve les mesures. Donc  $(X, \mathcal{B}, \mu, \phi)$  n'est pas totalement ergodique.

Réciproquement, supposons que  $(X, \mathcal{B}, \mu, \phi)$  ne soit pas totalement ergodique. Si déjà  $\phi$  n'est pas ergodique, alors il existe une partie mesurable  $\phi$ -invariante  $A$  de  $X$  de mesure ni pleine ni nulle. Notons  $\mu'$  la mesure sur l'espace discret  $\{0, 1\}$  définie par  $\mu'(\{1\}) = \mu(A)$  et  $\mu'(\{0\}) = \mu({}^c A)$ . Alors la fonction caractéristique  $\mathbb{1}_A : X \rightarrow X' = \{0, 1\}$  de  $A$  définie par  $x \mapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$  est une semi-conjugaison de systèmes dynamiques mesurés entre  $(X, \mu, \phi)$  et  $(\{0, 1\}, \mu', \text{id}_{\{0,1\}})$ .

Sinon, soit  $n \in \mathbb{N} \setminus \{0, 1\}$  le plus petit entier tel que  $\phi^n$  ne soit pas ergodique. Considérons une partie  $\phi^n$ -invariante  $A$  de  $X$  de mesure ni pleine ni nulle. Posons  $f = \sum_{i=0}^{n-1} \mathbb{1}_A \circ \phi^i$ , qui est une fonction mesurable à valeurs entières, d'intégrale  $\int_X f d\mu = n\mu(A)$  par l'invariance de  $\mu$ . Comme  $\mathbb{1}_A \circ \phi^k = \mathbb{1}_A$ , la fonction  $f$  est  $\phi$ -invariante, donc  $\mu$ -presque partout égal à une constante entière  $k$  par l'ergodicité de  $\phi$ . Nous avons donc  $\mu(A) = \frac{k}{n}$  et  $k \in \llbracket 1, n-1 \rrbracket$ . Si  $k \geq 2$ , alors les parties  $\phi^{-i}(A)$  pour  $i \in \llbracket 0, n-1 \rrbracket$ , qui ont la même mesure, ne sont pas deux à deux disjointes modulo mesure nulle. Soit  $i \in \llbracket 0, n-1 \rrbracket$  tel que la mesure de  $A' = A \cap \phi^{-i}(A)$  soit strictement positive. Notons que  $A'$  est  $\phi^n$ -invariante et n'est pas presque partout  $\phi^i$ -invariante car  $\phi^i$  est ergodique par la minimalité de  $n$ . Donc  $\mu(A') < \mu(A)$  et en itérant, nous pouvons donc supposer que  $\mu(A) = \frac{1}{n}$ . Remarquons que  $X = \bigcup_{i=1}^{n-1} \phi^{-i}(A)$  par l'ergodicité de  $f$  et puisque  $A$  est  $\phi^n$ -invariante. Cette réunion est donc presque partout disjointe, et quitte à modifier par des ensembles de mesure nulle, nous pouvons la supposer disjointe. Posons alors  $X' = \mathbb{Z}/n\mathbb{Z}$ ,  $\mu'$  l'équiprobabilité sur  $X'$  et  $\phi' : x \rightarrow x+1 \pmod n$  comme dans la solution de la question (1). L'application  $h : X \rightarrow X'$  valant  $i$  sur  $\phi^{-i}(A)$  pour tout  $i \in \llbracket 0, n-1 \rrbracket$  est alors une semi-conjugaison entre  $(X, \mathcal{B}, \mu, \phi)$  et un facteur fini non trivial de  $(X, \mathcal{B}, \mu, \phi)$ .

**Correction de l'exercice E.18.** (1) Comme  $\phi$  est ergodique et puisque les exposants de Lyapounov de  $(M_n)_{n \in \mathbb{N}}$  sont mesurables et presque partout invariants, ils sont presque partout constants, et donc presque partout égaux à leur intégrale sur  $X$ .

(2) Par les formules d'Euler  $\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$  et  $\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$ , nous avons immédiatement  $R_x = e^{-2i\pi x} \tilde{R}(e^{2i\pi x})$ .

(3) Pour tout  $z \in \mathbb{C}$ , considérons l'application  $C_n : \mathbb{C} \rightarrow \mathcal{M}_2(\mathbb{C})$  définie par

$$z \mapsto H\tilde{R}(e^{2\pi i(n-1)\alpha} z)H\tilde{R}(e^{2\pi i(n-2)\alpha} z) \dots H\tilde{R}(e^{2\pi i\alpha} z)H\tilde{R}(z).$$

Notons que  $C_n$  est une application holomorphe (car toute somme finie de produits finis de fonctions holomorphes est holomorphe). Donc l'application  $z \mapsto \ln \|C_n\|_\infty$ , qui est le minimum de quatre logarithmes de valeur absolue de fonctions holomorphes, est sous-harmonique. Nous avons  $\lambda_{\max} = \lim_{n \rightarrow +\infty} \frac{1}{n} \int_0^1 f_n(e^{2i\pi x}) dx$  presque partout par les questions (1) et (2).



(4) Par la définition d'une fonction sous-harmonique et du rayon spectral  $\rho(A)$  d'une matrice  $A$ , nous avons

$$\lambda_{\max} \geq \lim_{n \rightarrow +\infty} \frac{1}{n} f_n(0) = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|(H\tilde{R}(0))^n\|_{\infty} = \ln \rho(H\tilde{R}(0)).$$

Un calcul élémentaire montre que les valeurs propres de  $H\tilde{R}(0) = \frac{1}{2} \begin{pmatrix} c & -c^i \\ c^{-1} & c^{-1} \end{pmatrix}$  sont 0 et  $\frac{c+c^{-1}}{2}$ , et le résultat en découle.

### 3 Mélange et vitesses de mélange

#### 3.1 Mélange fort

Nous introduisons dans cette partie une propriété plus forte que l'ergodicité.

Si  $(X, \mathcal{B}, \mu)$  est un espace de probabilité, nous notons

$$\mathbb{L}_0^2(X, \mu) = \left\{ f \in \mathbb{L}^2(X, \mu) : \int_X f d\mu = 0 \right\} .$$

qui est le sous-espace de Hilbert (en tant qu'hyperplan fermé) de  $\mathbb{L}^2(X, \mu)$  orthogonal à la droite vectorielle des fonctions (presque partout) constantes. Dans l'énoncé ci-dessous, le temps  $t$  varie dans  $\mathbb{N}$ ,  $\mathbb{Z}$  ou  $\mathbb{R}$  en fonction du type de système dynamique.

**Proposition 3.1.** *Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé. Les conditions suivantes sont équivalentes.*<sup>87</sup>

(1) *Le système dynamique  $(X, \mathcal{B}, \mu, \phi)$  est mélangeant, c'est-à-dire*

$$\forall A, B \in \mathcal{B}, \quad \lim_{t \rightarrow +\infty} \mu(A \cap \phi^{-t}(B)) = \mu(A) \mu(B) .$$

(2) *Il existe des  $\pi$ -systèmes  $\mathcal{E}$  et  $\mathcal{E}'$  de  $\sigma$ -algèbre engendrée égale à  $\mathcal{B}$ , tels que*

$$\forall A \in \mathcal{E}, \forall B \in \mathcal{E}', \quad \lim_{t \rightarrow +\infty} \mu(A \cap \phi^{-t}(B)) = \mu(A) \mu(B) .$$

(3) *Nous avons*

$$\forall f, g \in \mathbb{L}^2(X, \mu), \quad \lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g d\mu = \left( \int_X f d\mu \right) \left( \int_X g d\mu \right) .$$

(4) *Il existe des parties  $E$  et  $E'$  de  $\mathbb{L}^2(X, \mu)$ , engendrant des sous-espaces vectoriels denses dans  $\mathbb{L}^2(X, \mu)$ , telles que*

$$\forall f \in E, \forall g \in E', \quad \lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g d\mu = \left( \int_X f d\mu \right) \left( \int_X g d\mu \right) .$$

(5) *Nous avons*

$$\forall f, g \in \mathbb{L}_0^2(X, \mu), \quad \lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g d\mu = 0 .$$

(6) *Il existe des parties  $E$  et  $E'$  de  $\mathbb{L}_0^2(X, \mu)$ , engendrant des sous-espaces vectoriels denses dans  $\mathbb{L}_0^2(X, \mu)$ , telles que*

$$\forall f \in E, \forall g \in E', \quad \lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g d\mu = 0 .$$

**Remarques.** (i) Lorsque la mesure  $\mu$  est sous-entendue, nous dirons aussi que la transformation ou le flot  $\phi$  est *mélangeante*. Lorsque la transformation ou le flot  $\phi$  est sous-entendue, nous dirons aussi que la mesure  $\mu$  est *mélangeante*. Certains auteurs disent *mélange fort*

<sup>87</sup>. Toujours avec la convention que  $t \in \mathbb{N}$  lorsque  $(X, \mathcal{B}, \mu, \phi)$  est à temps discret

au lieu de mélange, le *mélange faible* signifiant que l'action diagonale de  $\phi$  sur  $X \times X$  est ergodique pour  $\mu \times \mu$ .

(ii) La propriété de mélange signifie que, pour les grands temps  $t$ , les évènements  $A$  et  $\phi^{-t}(B)$  sont asymptotiquement indépendants.

(iii) Un système dynamique probabilisé mélangeant est ergodique. En effet, lorsque le temps est discret (la démonstration est analogue en temps continu), si  $A = \phi^{-1}(A)$  est une partie mesurable, alors nous avons  $\mu(A) = \mu(A \cap \phi^{-n}(A))$  pour tout  $n \in \mathbb{N}$ , et  $\lim_{n \rightarrow +\infty} \mu(A \cap \phi^{-n}(A)) = \mu(A)^2$ . Donc  $\mu(A) = \mu(A)^2$ , et  $\mu(A)$  vaut 0 ou 1.

(iv) Notons que par l'invariance de la mesure et par symétrie, si le système dynamique  $(X, \mathcal{B}, \mu, \phi)$  est à temps discret et inversible, ou s'il est à temps continu, alors il est mélangeant si et seulement si son inverse l'est, c'est-à-dire

$$\lim_{t \rightarrow +\infty} \int_X (f \circ \phi^{-t}) g d\mu = \left( \int_X f d\mu \right) \left( \int_X g d\mu \right).$$

pour tous les  $f, g$  dans  $\mathbb{L}^2(X, \mu)$  (ou dans des parties engendrant des sous-espaces vectoriels denses).

(v) Il est par exemple possible de prendre pour  $E = E'$  une base hilbertienne de  $\mathbb{L}^2(X, \mu)$  dans l'assertion (4) ou de  $\mathbb{L}_0^2(X, \mu)$  dans l'assertion (6). Si  $\mathbb{L}^2(X, \mu)$  est séparable,<sup>88</sup> alors les bases hilbertiennes sont dénombrables, et cela réduit la démonstration de la propriété de mélange à un ensemble dénombrable de vérifications.

(vi) L'espace  $\mathbb{L}^2(X, \mu)$  a le gros avantage d'être un espace de Hilbert. Mais de très nombreux sous-espaces vectoriels denses sont parfois utiles, en choisissant celui approprié en fonction du problème à étudier. Voici quelques exemples (rappelons que nous supposons que  $\mu$  est une mesure de probabilité).

- Le sous-espace vectoriel des *fonctions étagées*, c'est-à-dire les combinaisons linéaires finies de fonctions indicatrices de parties mesurables de  $X$ , est dense dans  $\mathbb{L}^2(X, \mu)$ .

- Plus généralement, si  $\mathcal{E}$  est un  $\pi$ -système de  $\sigma$ -algèbre engendrée égale à  $\mathcal{B}$ , alors le sous-espace vectoriel des combinaisons linéaires finies de fonctions indicatrices d'éléments de  $\mathcal{E}$  est dense dans  $\mathbb{L}^2(X, \mu)$ .

- Si  $p \in [2, +\infty]$ , l'espace vectoriel  $\mathbb{L}^p(X, \mu)$  est contenu et dense dans  $\mathbb{L}^2(X, \mu)$ . Si  $p \in [1, 2]$ , l'espace vectoriel  $\mathbb{L}^p(X, \mu) \cap \mathbb{L}^2(X, \mu)$  est dense dans  $\mathbb{L}^2(X, \mu)$ . En effet, les fonctions étagées sont dans  $\mathbb{L}^p(X, \mu)$  pour tout  $p \in [1, +\infty]$ .

- Si  $X$  est un espace topologique localement compact muni de sa tribu borélienne, alors les espaces vectoriels  $C_0^0(X)$  des fonctions (à valeurs complexes) continues nulles à l'infini (voir la partie 0.3 pour une définition) et  $C_c^0(X)$  des fonctions (à valeurs complexes) continues à support compact sont denses dans  $\mathbb{L}^2(X, \mu)$ .

- Si  $X$  est une variété différentielle munie de sa tribu borélienne, alors pour tout  $k \in \mathbb{N} \cup \{\infty\}$ , l'espace vectoriel  $C_c^k(X)$  des fonctions (à valeurs complexes) différentiables de classe  $C^k$  à support compact est dense dans  $\mathbb{L}^2(X, \mu)$ .

**Démonstration de la proposition 3.1.** Rappelons que par l'inégalité de Cauchy-Schwarz, le produit de deux fonctions  $\mathbb{L}^2$  est  $\mathbb{L}^1$ . Puisque la mesure est finie, une fonction  $\mathbb{L}^2$  est  $\mathbb{L}^1$ . Donc les différents termes sont bien définis.

<sup>88.</sup> Comme  $\mu$  est une mesure de probabilité, il suffit pour cela que la  $\sigma$ -algèbre de  $X$  soit engendrée par une partie dénombrable (par exemple si  $X$  est métrisable séparable), voir par exemple [Coh].

Montrons que (5)  $\Leftrightarrow$  (3)  $\Rightarrow$  (1)  $\Rightarrow$  (2)  $\Rightarrow$  (4)  $\Leftrightarrow$  (6)  $\Rightarrow$  (5), ce qui montrera le résultat.

Notons  $P(f, g)$  la propriété  $\lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g d\mu = (\int_X f d\mu)(\int_X g d\mu)$ . Elle est vérifiée si  $f$  ou si  $g$  est constante (car  $\phi$  préserve  $\mu$ ). Par bilinéarité, les assertions (3) et (5), ainsi que les assertions (4) et (6), sont donc équivalentes.

L'assertion (3) implique l'assertion (1), car il suffit de prendre les fonctions indicatrices  $f = \mathbb{1}_B$  et  $g = \mathbb{1}_A$ . L'assertion (1) implique immédiatement l'assertion (2). Puisque l'espace vectoriel des combinaisons linéaires de fonctions indicatrices d'éléments de  $\mathcal{E}$  est dense dans  $\mathbb{L}^2(X, \mu)$ , l'assertion (2) implique l'assertion (4). Il suffit maintenant de démontrer que l'assertion (6) implique l'assertion (5) pour terminer la démonstration de la proposition 3.1.

Pour tout  $g$  dans  $\mathbb{L}_0^2(X, \mu)$ , notons  $H_g$  l'ensemble des fonctions  $f$  dans  $\mathbb{L}_0^2(X, \mu)$  telles que la propriété  $P(f, g)$  soit vraie, qui est un sous-espace vectoriel de  $\mathbb{L}_0^2(X, \mu)$  par linéarité, et montrons qu'il est fermé dans  $\mathbb{L}_0^2(X, \mu)$ .

Nous pouvons supposer que  $g$  n'est pas presque partout nulle. Soit  $f$  dans  $\overline{H_g}$ . Pour tout  $\epsilon > 0$ , il existe un élément  $f_1$  dans  $H_g$  tel que  $\|f - f_1\|_{\mathbb{L}^2} \leq \frac{\epsilon}{2\|g\|_{\mathbb{L}^2}}$ . Nous avons alors  $|\int_X (f \circ \phi^t) g d\mu| \leq T_1 + T_2$  avec

$$T_1 = \left| \int_X ((f - f_1) \circ \phi^t) g d\mu \right| \leq \|f - f_1\|_{\mathbb{L}^2} \|g\|_{\mathbb{L}^2} \leq \epsilon/2$$

pour tout temps  $t$ , par l'inégalité de Cauchy-Schwarz et l'invariance de  $\mu$ , et

$$T_2 = \left| \int_X (f_1 \circ \phi^t) g d\mu \right| \leq \epsilon/2$$

pour tout temps  $t$  assez grand car  $f_1 \in H_g$  et  $\int_X g d\mu = 0$ . Ceci montre que nous avons  $\lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g d\mu = 0$ . Donc  $f$  appartient à  $H_g$ , et  $H_g$  est fermé.

Un raisonnement analogue montre que pour tout  $f$  dans  $\mathbb{L}_0^2(X, \mu)$ , le sous-espace vectoriel des fonctions  $g$  dans  $\mathbb{L}_0^2(X, \mu)$ , telles que la propriété  $P(f, g)$  est vraie, est fermé dans  $\mathbb{L}_0^2(X, \mu)$ . Puisqu'une partie fermée et dense d'un espace topologique est égal à cet espace, il est alors clair que l'assertion (6) implique l'assertion (5).  $\square$

**Proposition 3.2.** *Soit  $(\Omega, \mathcal{B}, \mu)$  un espace de probabilité. Alors les décalages à gauche  $\sigma : \Omega^{\mathbb{Z}} \rightarrow \Omega^{\mathbb{Z}}$  et  $\sigma_+ : \Omega^{\mathbb{N}} \rightarrow \Omega^{\mathbb{N}}$  sont mélangeants (donc ergodiques) pour les mesures de Bernoulli  $\mu^{\mathbb{Z}}$  et  $\mu^{\mathbb{N}}$  respectivement.*

**Démonstration.** Traitons le cas unilatère, le cas bilatère est similaire. Le décalage  $\sigma_+$  laisse invariante la mesure produit  $\mu^{\mathbb{N}}$  par l'exemple (2) de la partie 0.2. Comme les fonctions caractéristiques des cylindres engendrent un sous-espace vectoriel dense de l'espace  $\mathbb{L}^2(\Omega^{\mathbb{N}}, \mu^{\mathbb{N}})$ , en utilisant la proposition 3.1 (4), il suffit de montrer la propriété de mélange

$$\lim_{n \rightarrow +\infty} \mu^{\mathbb{N}}(A \cap \sigma_+^{-n}(B)) = \mu^{\mathbb{N}}(A) \mu^{\mathbb{N}}(B)$$

pour les cylindres  $A$  et  $B$ . Or si  $A = [\omega_0 \in E_0, \dots, \omega_k \in E_k]$  et  $B = [\omega_0 \in F_0, \dots, \omega_\ell \in F_\ell]$ , alors pour tout  $n > k$ , l'intersection  $A \cap \sigma_+^{-n}(B)$  est le cylindre

$$\{(\omega_i)_{i \in \mathbb{N}} : \omega_0 \in E_0, \dots, \omega_k \in E_k, \omega_n \in F_0, \dots, \omega_{n+\ell} \in F_\ell\}.$$

Donc  $\mu^{\mathbb{N}}(A \cap \sigma_+^{-n}(B)) = \mu(A) \mu(B)$  par la définition de la mesure produit  $\mu^{\mathbb{N}}$ .  $\square$

### 3.2 Mélange exponentiel

Voici une propriété plus forte que celle du mélange. Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé. Appelons un *coefficient de corrélation* de  $(X, \mathcal{B}, \mu, \phi)$  au temps  $t$  toute quantité

$$c_t(f, g) = c_{t, \phi}(f, g) = \int_X (f \circ \phi^t) g \, d\mu - \left( \int_X f \, d\mu \right) \left( \int_X g \, d\mu \right),$$

où  $f, g \in \mathbb{L}^2(X)$ . Notons que l'application  $(f, g) \mapsto c_t(f, g)$  est une forme bilinéaire continue (avec  $|c_t(f, g)| \leq 2\|f\|_{\mathbb{L}^2} \|g\|_{\mathbb{L}^2}$ ) sur  $\mathbb{L}^2(X)$ , telle que si  $\phi'$  est une transformation préservant la mesure et commutant avec  $\phi$ , alors

$$c_t(f \circ \phi', g \circ \phi') = c_t(f, g).$$

L'étude des propriétés d'annulation asymptotique effectives, c'est-à-dire avec un contrôle de la vitesse de convergence vers 0 quand  $t \rightarrow +\infty$ , des coefficients de corrélation est actuellement en plein développement (voir par exemple [Pol, Rat, You, Sar3, BaL, BurMMW, Tsu, Mel, Bal, Nau]).

Soit  $E$  un sous-espace vectoriel dense de  $\mathbb{L}^2(X, \mu)$ , muni d'une norme  $\|\cdot\|_E$  (qui n'est pas forcément la norme  $\mathbb{L}^2$  elle-même). Soit  $\psi : \mathbb{N} \rightarrow ]0, +\infty[$  si  $(X, \mathcal{B}, \mu, \phi)$  est à temps discret, et  $\psi : \mathbb{R} \rightarrow ]0, +\infty[$  si  $(X, \mathcal{B}, \mu, \phi)$  est à temps continu, une application telle que  $\lim_{t \rightarrow +\infty} \psi(t) = 0$ .

Nous dirons que le système dynamique mesuré  $(X, \mathcal{B}, \mu, \phi)$  est  *$\psi$ -mélangeant sur  $E$* , ou qu'il est à  *$\psi$ -décroissance des corrélations sur  $E$*  s'il existe  $c > 0$  tel que pour tout temps  $t \geq 0$ , nous avons, pour tous les  $f, g \in E$ ,

$$\left| \int_X (f \circ \phi^t) g \, d\mu - \left( \int_X f \, d\mu \right) \left( \int_X g \, d\mu \right) \right| \leq c \|f\|_E \|g\|_E \psi(t).$$

Notons qu'être  $\psi$ -mélangeant sur un sous-espace vectoriel dense  $E$  implique être mélangeant, par l'équivalence des définitions dans la proposition 3.1, car  $\lim_{t \rightarrow +\infty} \psi(t) = 0$ . Remarquons que si  $(X, \mathcal{B}, \mu, \phi)$  est à temps continu, alors la formule centrée ci-dessus est alors valable pour tout  $t \in \mathbb{R}$  par l'invariance de la mesure et par symétrie.

**Exemples.** S'il existe  $\kappa > 0$  tel que le système dynamique mesuré  $(X, \mathcal{B}, \mu, \phi)$  soit  $\psi$ -mélangeant sur  $E$  avec  $\psi(t) = e^{-\kappa t}$ , nous dirons que  $(X, \mathcal{B}, \mu, \phi)$  est *exponentiellement mélangeant sur  $E$*  ou à *décroissance exponentielle des corrélations sur  $E$* .

S'il existe un entier  $k \in \mathbb{N} - \{0\}$  tel que le système dynamique mesuré  $(X, \mathcal{B}, \mu, \phi)$  soit  $\psi$ -mélangeant sur  $E$  avec  $\psi(t) = (|t| + 1)^{-k}$ , nous dirons que  $(X, \mathcal{B}, \mu, \phi)$  est *polynomialement mélangeant sur  $E$*  ou à *décroissance polynomiale des corrélations sur  $E$* .

Nous dirons que  $(X, \mathcal{B}, \mu, \phi)$  est *rapidement mélangeant sur  $E$*  ou à *décroissance rapide des corrélations sur  $E$*  si pour tout  $k \in \mathbb{N} - \{0\}$ , le système dynamique mesuré  $(X, \mathcal{B}, \mu, \phi)$  est  $\psi$ -mélangeant sur  $E$  avec  $\psi(t) = (|t| + 1)^{-k}$ .

Les sous-espaces vectoriels denses  $E$  de  $\mathbb{L}^2(X, \mu)$  les plus employés sont les suivants.

Soit  $\alpha \in ]0, 1]$ . Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé, où  $X$  est un espace métrique muni de sa tribu borélienne  $\mathcal{B}$ . Notons  $C_b^\alpha(X)$  le sous-espace vectoriel (dense)

de  $\mathbb{L}^2(X, \mu)$  des fonctions continues bornées  $\alpha$ -höldériennes<sup>89</sup> sur  $X$ , muni de la norme  $\alpha$ -höldérienne  $f \mapsto \|f\|_{\infty, \alpha}$ . Notons que l'espace vectoriel normé  $(C_b^\alpha(X), \|\cdot\|_{\infty, \alpha})$  est un espace de Banach.

Soient  $p \in [2, +\infty]$  et  $k \in \mathbb{N}$ . Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé, où  $X$  est un ouvert de  $\mathbb{R}^N$  muni de sa tribu borélienne.<sup>90</sup> Notons  $E_{p,k}$  le sous-espace vectoriel (dense) de  $\mathbb{L}^2(X, \mu)$  des fonctions différentiables  $f$  de classe  $C^k$ , qui sont dans  $\mathbb{L}^p(X, \mu)$  et dont toutes les dérivées partielles jusqu'à l'ordre  $k$  sont dans  $\mathbb{L}^p(X, \mu)$ , muni de la *norme*<sup>91</sup> *de Sobolev*<sup>92</sup>, notée  $\|\cdot\|_{p,k}$  ou  $\|\cdot\|_{W^{p,k}}$ , définie par

$$\|f\|_{p,k} = \|f\|_p + \sum_{\substack{1 \leq \ell \leq k \\ 1 \leq i_1, \dots, i_\ell \leq N}} \|\partial_{i_1, \dots, i_\ell} f\|_p.$$

### 3.3 Un peu de théorie spectrale de l'opérateur de Koopman

Nous renvoyons par exemple à [Pau6, Rud2] pour une introduction générale à la théorie spectrale des opérateurs bornés des espaces de Hilbert, mais nous rappelons tout ce dont nous aurons besoin ci-dessous.

Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret. Notons pour simplifier  $\mathbb{L}^2(\mu) = \mathbb{L}^2(X, \mathcal{B}, \mu; \mathbb{C})$ , et notons

$$\langle \cdot, \cdot \rangle : (f, g) \mapsto \int_X f \bar{g} \, d\mu$$

le produit scalaire (hermitien) de  $\mathbb{L}^2(\mu)$ .

Considérons l'*opérateur de Koopman*

$$U_\phi : \mathbb{L}^2(\mu) \rightarrow \mathbb{L}^2(\mu) \quad \text{défini par} \quad f \mapsto f \circ \phi.$$

---

89. Pour tout  $\alpha \in ]0, 1]$ , si  $X$  et  $Y$  sont des espaces métriques, une application continue  $f : X \rightarrow Y$  est dite  $\alpha$ -höldérienne s'il existe  $c > 0$  tel que pour tous les  $x, y \in X$ , si  $d(x, y) \leq 1$ , alors

$$d(f(x), f(y)) \leq c d(x, y)^\alpha.$$

L'espace vectoriel  $C_b^\alpha(X)$  des applications continues bornées  $\alpha$ -höldériennes d'un espace métrique  $X$  dans  $\mathbb{R}$ , muni de la *norme höldérienne*

$$\|f\|_{\infty, \alpha} = \|f\|_\infty + \sup_{x, y \in X, 0 < d(x, y) \leq 1} \frac{|f(x) - f(y)|}{d(x, y)^\alpha},$$

est un espace de Banach.

90. Ceci se généralise à toutes les variétés différentielles. Nous renvoyons à la partie 4 pour le cas particulier du tore  $\mathbb{T}^N = \mathbb{R}^N / \mathbb{Z}^N$ , simplifié par l'identification entre fonctions sur  $\mathbb{T}^N$  et fonctions  $\mathbb{Z}^N$ -périodiques sur  $\mathbb{R}^N$ .

91. Attention, cette norme sur  $E_{p,k}$  n'est pas complète. Le complété de  $E_{p,k}$  pour cette norme est noté  $W^{p,k}(X)$ .



92. Serguei Sobolev  
1908-1989



Joseph Fourier  
1768-1830



Salomon Bochner  
1899-1982

**Remarques.** (i) L'opérateur de Koopman  $U_\phi$  est un opérateur linéaire préservant le produit scalaire<sup>93</sup> (car  $\phi$  préserve  $\mu$ ) de l'espace de Hilbert complexe  $\mathbb{L}^2(\mu)$ , donc continu avec norme d'opérateur  $\|U_\phi\| = 1$ .

(ii) L'hyperplan fermé  $\mathbb{L}^2(\mu) = \mathbb{L}_0^2(X, \mathcal{B}, \mu)$  de  $\mathbb{L}^2(\mu)$  (l'orthogonal à la droite vectorielle des fonctions (presque partout) constantes sur  $X$  défini au début de la partie 3.1), ainsi que sa droite vectorielle orthogonale, sont invariants par  $U_\phi$ .

(iii) Si le système dynamique  $(X, \mathcal{B}, \mu, \phi)$  est inversible, alors l'opérateur  $U_\phi$  est inversible et l'opérateur adjoint  $(U_\phi)^*$  de  $U_\phi$  est, par l'invariance de  $\mu$  par  $\phi$ ,

$$(U_\phi)^* = U_{\phi^{-1}} = U_\phi^{-1} .$$

Le calcul de l'adjoint de  $U_\phi$  dans le cas non inversible est souvent intéressant et utile (voir par exemple l'exercice E.26), et a été développé dans les théories des opérateurs de transfert (voir par exemple les travaux de Ruelle et [Bal]).

(iv) Si  $(X', \mathcal{B}', \mu', \phi')$  est un système dynamique probabilisé à temps discret conjugué à  $(X, \mathcal{B}, \mu, \phi)$ , alors les opérateurs de Koopman  $U_\phi$  et  $U_{\phi'}$  sont unitairement conjugués, c'est-à-dire qu'il existe un isomorphisme d'espaces de Hilbert complexes  $u : \mathbb{L}^2(\mu') \rightarrow \mathbb{L}^2(\mu)$  tel que  $U_{\phi'} = u \circ U_\phi \circ u^{-1}$ . En effet, soit  $h : X \rightarrow X'$  une conjugaison mesurée (quitte à enlever des parties invariantes mesurables de mesure nulle). Alors l'application  $U_h : \mathbb{L}^2(\mu') \rightarrow \mathbb{L}^2(\mu)$  définie par  $f \mapsto f \circ h$  est un isomorphisme linéaire (d'inverse  $(U_h)^{-1} = U_{h^{-1}}$ ) qui est un isomorphisme d'espaces de Hilbert (car  $h_*\mu = \mu'$ , donc  $\int_X (f \circ h) \overline{(g \circ h)} d\mu = \int_{X'} f \overline{g} d\mu'$  pour tous les  $f, g \in \mathbb{L}^2(\mu')$ ). De plus  $U_{\phi'} = U_h \circ U_\phi \circ U_h^{-1}$ , car  $f \circ \phi' = ((f \circ h) \circ \phi) \circ h^{-1}$  pour tout  $f \in \mathbb{L}^2(\mu')$ .

En particulier, tout invariant de conjugaison de  $U_\phi$  (comme son spectre ou son spectre ponctuel, voir ci-dessous, et on parle alors d'invariant spectraux de  $(X, \mathcal{B}, \mu, \phi)$ ) est un invariant de conjugaison mesurée de systèmes dynamiques probabilisés à temps discret.

Rappelons que

- un nombre complexe  $\lambda \in \mathbb{C}$  est une valeur propre d'un opérateur linéaire continu  $u \in \mathcal{L}(\mathcal{H})$  d'un espace de Hilbert complexe  $\mathcal{H}$  si et seulement si le sous-espace vectoriel  $\ker(u - \lambda \text{id})$  de  $\mathcal{H}$  (dont les éléments sont appelés les vecteurs propres<sup>94</sup> de  $\mathcal{H}$  pour la valeur propre  $\lambda$ ) est non nul,
- le spectre ponctuel<sup>95</sup>  $\text{Vp}(u)$  de  $u$  est l'ensemble des valeurs propres de  $u$ ,
- si  $u$  préserve le produit scalaire, alors  $\text{Vp}(u)$  est contenu dans  $\mathbb{S}_1$ , car si  $f$  est un vecteur propre non nul de  $u$  pour la valeur propre  $\lambda$ , alors  $|\lambda|^2 = \frac{\langle u(f), u(f) \rangle}{\langle f, f \rangle} = 1$ .

93. c'est-à-dire que pour tous les  $f, g \in \mathbb{L}^2(\mu)$ , nous avons  $\langle U_\phi f, U_\phi g \rangle = \int_X (f \circ \phi) \overline{(g \circ \phi)} d\mu = \langle f, g \rangle$ .

94. Certains ouvrages supposent dans leur définition que les vecteurs propres sont non nuls, mais il est préférable de dire « vecteur propre non nul » en cas de besoin.

95. Il est important de ne pas confondre en dimension infinie le spectre ponctuel  $\text{Vp}(u)$  avec le spectre  $\text{Sp}(u)$  de  $u$ , où  $\text{Sp}(u)$  est l'ensemble des  $\lambda \in \mathbb{C}$  tels que  $u - \lambda \text{id}$  ne soit pas inversible dans  $\mathcal{L}(\mathcal{H})$ , voir par exemple [Pau6]. Si  $u' \in \mathcal{L}(\mathcal{H}')$  est un opérateur linéaire continu d'un espace de Hilbert complexe  $\mathcal{H}'$  qui est conjugué à  $u$ , c'est-à-dire tel qu'il existe un isomorphisme linéaire continu (donc d'inverse continu par le théorème de Banach)  $v \in \mathcal{GL}(\mathcal{H}, \mathcal{H}')$  tel que  $u' = v \circ u \circ v^{-1}$ , alors

$$\text{Sp}(u') = \text{Sp}(u) \quad \text{et} \quad \text{Vp}(u') = \text{Vp}(u)$$

car pour tout  $\lambda \in \mathbb{C}$ , l'opérateur linéaire  $u' - \lambda \text{id} = v \circ (u - \lambda \text{id}) \circ v^{-1}$  est inversible ou injectif si et seulement si  $u - \lambda \text{id}$  l'est.

**Exemples.** (1) Nous avons  $\text{Vp}(U_\phi) \subset \mathbb{S}_1$  et 1 est valeur propre de  $U_\phi$ , car  $\ker(U_\phi - \text{id})$  contient la droite vectorielle des fonctions (presque partout) constantes.

(2) Soient  $\alpha \in \mathbb{R} - \mathbb{Q}$  et  $(X = \mathbb{S}_1, \mu = \mu_{\mathbb{S}_1}, \phi = R_\alpha : z \mapsto e^{2i\pi\alpha}z)$  le système de Kronecker de la rotation irrationnelle du cercle d'angle  $2\pi\alpha$ . Alors l'opérateur de Koopman  $U_\phi$  est diagonal dans la base hilbertienne  $(z \mapsto z^m)_{m \in \mathbb{Z}}$  de  $\mathbb{L}^2(\mu_{\mathbb{S}_1})$  :

$$\forall m \in \mathbb{Z}, \quad U_\phi(z^m) = e^{2i\pi m\alpha} z^m .$$

Donc  $\text{Vp}(U_\phi) = \{e^{2i\pi m\alpha} : m \in \mathbb{Z}\}$ .<sup>96</sup>

(3) L'assertion (4) avec  $p = 2$  de la liste des définitions équivalentes à l'ergodicité donnée dans la proposition 2.1 dit que l'ergodicité de  $\phi$  est équivalente au fait que la restriction de  $U_\phi$  à  $\mathbb{L}_0^2(\mu)$  n'admette pas 1 comme valeur propre.

**Exercice E.19.** Si  $\mathcal{I}$  est la tribu des invariants de  $(X, \mathcal{B}, \mu, \phi)$ , montrer que

$$\ker(U_\phi - \text{id}) = \mathbb{L}^2(X, \mathcal{I}, \mu) .$$

### Rappels sur la transformation de Fourier<sup>92</sup> des mesures positives sur le cercle

Pour toute mesure (borélienne) positive finie  $\mu$  sur le cercle  $\mathbb{S}_1$ , notons  $(c_m(\mu))_{m \in \mathbb{Z}}$  la transformée de Fourier de  $\mu$ , c'est-à-dire la famille indexée par  $\mathbb{Z}$  de ses coefficients de Fourier, où

$$\forall m \in \mathbb{Z}, \quad c_m(\mu) = \int_{z \in \mathbb{S}_1} z^{-m} d\mu(z) = \int_{e^{2i\pi\theta} \in \mathbb{S}_1} e^{-2i\pi m\theta} d\mu(e^{2i\pi\theta}) .$$

Par exemple, la transformée de Fourier de la mesure de Haar normalisée sur  $\mathbb{S}_1$ , définie par  $\mu_{\mathbb{S}_1} : f \mapsto \int_0^1 f(e^{2i\pi t}) dt$ , est

$$\forall m \in \mathbb{Z}, \quad c_m(\mu_{\mathbb{S}_1}) = \delta_{0,m} , \tag{30}$$

où  $\delta_{0,m}$  est le symbole de Kronecker, valant 1 si  $m = 0$  et 0 sinon.

Par exemple, la transformée de Fourier de la masse de Dirac unité  $\Delta_\lambda$  en  $\lambda \in \mathbb{S}_1$  est

$$\forall m \in \mathbb{Z}, \quad c_m(\Delta_\lambda) = \lambda^{-m} . \tag{31}$$

Rappelons des propriétés de base de la transformation de Fourier des mesures positives sur le cercle, dont une caractérisation fort utile (voir par exemple [Rud1][Die2, Chap. XIX]).

**Proposition 3.3.** (i) La transformation de Fourier  $\mathcal{F} : \mu \mapsto (c_m(\mu))_{m \in \mathbb{Z}}$  de l'espace  $\mathcal{M}_+(\mathbb{S}_1)$  des mesures positives finies sur  $\mathbb{S}_1$ , muni de la topologie faible-étoile<sup>97</sup>, à valeur dans l'espace topologique produit  $\mathbb{C}^{\mathbb{Z}}$  est continue injective, et un homéomorphisme sur son image. De plus, pour tous les  $\mu, \nu \in \mathcal{M}_+(\mathbb{S}_1)$ , nous avons

$$c_0(\mu) = \|\mu\|$$

96. Notons que le spectre  $\text{Sp}(U_\phi)$  de l'opérateur linéaire continu  $U_\phi$ , étant fermé (voir par exemple [Pau6, Théo. 1.25]), contenu dans le cercle car  $U_\phi$  est inversible unitaire (voir par exemple [Pau6, Exer. E.16]), et contenant la partie dense  $\text{Vp}(U_\phi)$  du cercle, est alors égal à tout le cercle :

$$\text{Vp}(U_{R_\alpha}) = \mathbb{S}_1 .$$

97. ou vague, voir la partie 0.3



et la transformation de Fourier des mesures positives est antisymétrique et positivement linéaire :

$$\forall m \in \mathbb{Z}, \quad \overline{c_m(\mu)} = c_{-m}(\mu) \quad \text{et} \quad \forall \lambda \in [0, +\infty[, \quad c_m(\lambda\mu + \nu) = \lambda c_m(\mu) + c_m(\nu).$$

En particulier, deux mesures positives finies  $\mu$  et  $\nu$  sur  $\mathbb{S}_1$  sont égales si et seulement si elles ont même masse totale et si  $c_m(\mu) = c_m(\nu)$  pour tout  $m \in \mathbb{N} \setminus \{0\}$ .

(ii) (**Théorème de Bochner**<sup>92</sup>) Une suite complexe  $(a_m)_{m \in \mathbb{Z}}$  est la transformée de Fourier d'une mesure positive finie sur  $\mathbb{S}_1$  si et seulement si elle est de type positif, c'est-à-dire si elle vérifie, pour toute partie finie  $E$  de  $\mathbb{Z}$  et toute suite finie  $(z_i)_{i \in E}$  dans  $\mathbb{C}$ , que

$$\sum_{i, j \in E} a_{i-j} z_i \overline{z_j} \geq 0.$$

**Démonstration.** (i) L'application  $\mathcal{F}$  est bien définie, et elle est clairement antisymétrique (ceci utilise le fait que la mesure soit positive, donc réelle) et positivement linéaire. Comme le  $m$ -ème coefficient de Fourier est l'évaluation des mesures positives sur la fonction continue  $z \mapsto z^m$  sur le cercle, et par la définition de la topologie produit, l'application  $\mathcal{F}$  est continue. Comme l'espace vectoriel engendré par les fonctions  $z \mapsto z^m$  pour  $m \in \mathbb{Z}$  est dense dans  $C^0(\mathbb{S}_1; \mathbb{C})$  par le théorème de Stone-Weierstrass (nous y reviendrons, voir la note de bas de page 105), l'application  $\mathcal{F}$  est injective. Comme l'espace topologique produit  $\mathbb{C}^{\mathbb{Z}}$  est séparé (car métrisable), comme l'espace des mesures de probabilité sur le cercle est compact pour la topologie faible-étoile, et comme  $\mathcal{F}$  commute avec les homothéties de rapport positif, l'application  $\mathcal{F}$  est donc un homéomorphisme sur son image.

(ii) Pour tout  $n \geq 1$ , l'application  $P_n : z \mapsto \frac{1}{n} \sum_{1 \leq i, j \leq n} a_{i-j} z^{i-j}$  sur le cercle est continue et positive par l'hypothèse. Pour tout  $m \in \mathbb{Z}$  tel que  $m \leq n$ , le  $m$ -ème coefficient de Fourier de la mesure positive finie  $P_n \mu_{\mathbb{S}_1}$  est, par la linéarité de l'intégrale et par la formule (30), égal à

$$\begin{aligned} c_m(P_n \mu_{\mathbb{S}_1}) &= \int_{\mathbb{S}_1} z^{-m} \left( \frac{1}{n} \sum_{1 \leq i, j \leq n} a_{i-j} z^{i-j} \right) d\mu_{\mathbb{S}_1} = \frac{1}{n} \sum_{1 \leq i, j \leq n} a_{i-j} \int_{\mathbb{S}_1} z^{i-j-m} d\mu_{\mathbb{S}_1} \\ &= \frac{1}{n} \sum_{1 \leq i, j \leq n} a_{i-j} c_{m-i+j}(\mu_{\mathbb{S}_1}) = \frac{1}{n} \sum_{1 \leq i, j \leq n} a_{i-j} \delta_{0, m-i+j} \\ &= \frac{1}{n} \sum_{1 \leq i, j \leq n, i-j=m} a_{i-j} = \frac{n-m}{n} a_m \xrightarrow{n \rightarrow +\infty} a_m. \end{aligned}$$

En particulier, la masse totale de  $P_n \mu_{\mathbb{S}_1}$  est égale à  $a_0$ .

Toute valeur d'adhérence pour la topologie faible-étoile de la suite  $(P_n \mu_{\mathbb{S}_1})_{n \in \mathbb{N}}$  de mesures positives finies de masse totale constante  $a_0$ , qui existe par compacité (voir la proposition 0.8), est alors une mesure positive dont la transformée de Fourier est  $(a_m)_{m \in \mathbb{Z}}$  par la continuité de  $\mathcal{F}$ .  $\square$

Nous allons maintenant utiliser l'outil de la transformation de Fourier des mesures positives sur le cercle pour étudier des propriétés spectrales de l'opérateur de Koopman.

**Proposition 3.4.** Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret inversible.

- (1) Pour tout  $f \in \mathbb{L}^2(\mu)$ , il existe une et une seule mesure positive finie  $\sigma_f$  sur  $\mathbb{S}_1$ , appelée la mesure spectrale de  $f$ , telle que pour tout  $m \in \mathbb{Z}$ , nous ayons

$$c_m(\sigma_f) = \langle U_\phi^m f, f \rangle .$$

- (2) La masse totale de la mesure spectrale de  $f \in \mathbb{L}^2(\mu)$  est  $\|\sigma_f\| = \|f\|_{\mathbb{L}^2}^2$ , et en particulier, la mesure spectrale de  $f$  est une mesure de probabilité si et seulement si  $\|f\|_{\mathbb{L}^2} = 1$ .
- (3) La mesure spectrale est invariante par l'opérateur de Koopman :

$$\forall f \in \mathbb{L}^2(\mu), \quad \sigma_{U_\phi f} = \sigma_f .$$

- (4) Soient  $f \in \mathbb{L}^2(\mu)$  et  $\lambda \in \mathbb{C}$ . Nous avons

$$\sigma_{\lambda f} = |\lambda|^2 \sigma_f .$$

De plus, si  $\lambda \in \mathbb{S}_1$ , la fonction  $f$  est un vecteur propre de  $U_\phi$  pour la valeur propre  $\lambda$  si et seulement si  $\sigma_f = \|f\|_{\mathbb{L}^2}^2 \Delta_{\bar{\lambda}}$ .

- (5) Le système dynamique  $(X, \mathcal{B}, \mu, \phi)$  est mélangeant si et seulement si

$$\forall f \in \mathbb{L}_0^2(\mu), \quad \lim_{m \rightarrow +\infty} c_m(\sigma_f) = 0 .$$

En particulier, si le spectre ponctuel  $\text{Vp}(U_\phi|_{\mathbb{L}_0^2(\mu)})$  de la restriction de  $U_\phi$  à  $\mathbb{L}_0^2(\mu)$  est non vide, alors  $\phi$  n'est pas mélangeante. En effet, si  $f \in \mathbb{L}_0^2(\mu)$  est un vecteur propre unitaire de  $U_\phi$  de valeur propre  $\lambda$ , alors  $c_m(\sigma_f) = \langle U_\phi^m f, f \rangle = \lambda^m$  pour tout  $m \in \mathbb{N}$ , et comme  $|\lambda| = 1$ , la suite  $(c_m(\sigma_f))_{m \in \mathbb{N}}$  ne converge pas vers 0.

Par conséquent, par l'exemple (2) ci-dessus, une rotation irrationnelle du cercle n'est pas mélangeante.<sup>98</sup>

**Démonstration.** Notons que l'hypothèse d'inversibilité de  $\phi$  est nécessaire pour que l'opérateur de Koopman  $U_\phi$  soit inversible, en particulier afin de définir ses puissances négatives.

(1) Par le théorème de Bochner 3.3 (ii), l'existence de  $\sigma_f$  découle du fait que pour toute partie finie  $E$  de  $\mathbb{Z}$  et toute suite finie  $(z_i)_{i \in E}$  dans  $\mathbb{C}$ , par l'invariance par  $U_\phi$  du produit scalaire et par sesquilinearité, nous avons

$$\begin{aligned} \sum_{i,j \in E} \langle U_\phi^{i-j} f, f \rangle z_i \bar{z}_j &= \sum_{i,j \in E} z_i \langle U_\phi^i f, U_\phi^j f \rangle \bar{z}_j = \left\langle \sum_{i \in E} z_i U_\phi^i f, \sum_{j \in E} z_j U_\phi^j f \right\rangle \\ &= \left\| \sum_{k \in E} z_k U_\phi^k f \right\|^2 \geq 0 . \end{aligned}$$

L'unicité découle de l'injectivité de la transformation de Fourier des mesures positives sur  $\mathbb{S}_1$ .

98. Si  $\alpha \in \mathbb{R} - \mathbb{Q}$  et  $R_\alpha : z \mapsto e^{2i\pi\alpha} z$ , ceci se vérifie aussi de manière élémentaire en regardant les intervalles égaux  $A = B = \{e^{2i\pi t} : t \in [0, \frac{1}{2}]\}$  du cercle  $\mathbb{S}_1$ , car pour toute suite  $(n_k)_k \in \mathbb{N}$  telle que nous ayons  $\lim_{k \rightarrow +\infty} (n_k \alpha \bmod 1) = 0$  (qui existe car  $\alpha$  est irrationnel), nous avons

$$\lim_{k \rightarrow +\infty} \mu_{\mathbb{S}_1}(R_\alpha^{-n_k}(A) \cap B) = \frac{1}{2} \neq \frac{1}{4} = \mu_{\mathbb{S}_1}(A) \mu_{\mathbb{S}_1}(B) .$$

(2) Nous avons  $\|\sigma_f\| = c_0(\sigma_f) = \langle f, f \rangle = \|f\|_{\mathbb{L}^2}^2$ .

(3) Pour tous les  $m \in \mathbb{Z}$  et  $f \in \mathbb{L}^2(\mu)$ , puisque l'opérateur de Koopman commute avec ses puissances et préserve le produit scalaire, nous avons

$$c_m(\sigma_{U_\phi f}) = \langle U_\phi^m(U_\phi f), U_\phi f \rangle = \langle U_\phi^m f, f \rangle = c_m(\sigma_f).$$

L'injectivité de la transformation de Fourier dit donc que  $\sigma_{U_\phi f} = \sigma_f$ .

(4) La première affirmation est immédiate par la définition et l'injectivité de la transformation de Fourier car  $c_m(\sigma_{\lambda f}) = |\lambda|^2 c_m(\sigma_f) = c_m(|\lambda|^2 \sigma_f)$  pour tout  $m \in \mathbb{Z}$ . Pour montrer la seconde affirmation, nous pouvons donc supposer que  $\|f\| = 1$ .

Si  $f$  est un vecteur propre unitaire de  $U_\phi$  pour la valeur propre  $\lambda$ , alors pour tout  $m \in \mathbb{Z}$ , par la formule (31), nous avons

$$c_m(\sigma_f) = \langle U_\phi^m f, f \rangle = \langle \lambda^m f, f \rangle = \lambda^m = c_m(\Delta_{\bar{\lambda}}).$$

Le fait que  $\sigma_f = \Delta_{\bar{\lambda}}$  découle alors aussi de l'injectivité de la transformation de Fourier.

Réciproquement, si  $f$  est unitaire et  $\sigma_f = \Delta_{\bar{\lambda}}$ , alors  $\langle U_\phi f, f \rangle = c_1(\sigma_f) = c_1(\Delta_{\bar{\lambda}}) = \lambda$ . Donc, en rappelant que  $|\lambda| = 1$ , nous avons

$$\begin{aligned} \|U_\phi f - \lambda f\|^2 &= \|U_\phi f\|^2 - 2 \operatorname{Re} \langle U_\phi f, \lambda f \rangle + \|\lambda f\|^2 = 1 - 2 \operatorname{Re} (\bar{\lambda} \langle U_\phi f, f \rangle) + |\lambda|^2 \\ &= 2 - 2|\lambda|^2 = 0. \end{aligned}$$

Ceci montre que  $f$  est un vecteur propre de  $U_\phi$  pour la valeur propre  $\lambda$ .

(5) Comme  $c_m(\sigma_f) = \langle U_\phi^m f, f \rangle = \int_X f \circ \phi^m \bar{f} d\mu$ , l'assertion (5) découle de la proposition 3.1 (5) en prenant  $g = \bar{f}$ . Réciproquement, supposons que l'assertion (5) soit vérifiée. Pour tout  $f \in \mathbb{L}_0^2(\mu)$ , notons  $\mathcal{H}_f$  le plus petit sous-espace vectoriel fermé de l'espace de Hilbert complexe  $\mathbb{L}_0^2(\mu)$  contenant les fonctions  $U_\phi^m f$  pour  $m \in \mathbb{N}$ . L'ensemble  $\mathcal{H}'_f = \{g \in \mathbb{L}_0^2(\mu) : \lim_{n \rightarrow +\infty} \langle U_\phi^n f, g \rangle = 0\}$  est un sous-espace vectoriel fermé de  $\mathbb{L}_0^2(\mu)$ , qui contient la fonction  $f$  (par l'assertion (5)), et même tous les  $U_\phi^m f$  pour  $m \in \mathbb{N}$  par l'invariance du produit scalaire par  $U_\phi^m$  (qui commute avec les puissances de  $U_\phi$ ). Donc  $\mathcal{H}'_f$  contient  $\mathcal{H}_f$ . Par ailleurs, si  $g \in \mathbb{L}_0^2(\mu)$  est orthogonal à  $\mathcal{H}_f$ , alors  $\langle U_\phi^m f, g \rangle = 0$  pour tout  $m \in \mathbb{N}$ , et en particulier  $g \in \mathcal{H}'_f$ . Ainsi, le sous-espace vectoriel  $\mathcal{H}'_f$  contient à la fois  $\mathcal{H}_f$  et son orthogonal  $\mathcal{H}_f^\perp$  (qui est un supplémentaire de  $\mathcal{H}_f$  car  $\mathcal{H}_f$  est fermé). Par conséquent  $\mathcal{H}'_f = \mathbb{L}_0^2(\mu)$ , et  $\phi$  est mélangeante par la proposition 3.1 (5).  $\square$

**Exercice E.20.** Soient  $\mathcal{A}$  un alphabet fini de cardinal au moins 2 muni d'une mesure de probabilité  $\nu$ , et  $(X = \mathcal{A}^{\mathbb{Z}}, \mu = \nu^{\mathbb{Z}}, \phi = \sigma)$  le système de Bernoulli bilatère (avec sa mesure de Bernoulli) associé. Pour tout  $a \in \mathcal{A}$ , si  $[a]$  est le cylindre des suites  $(x_n)_{n \in \mathbb{Z}}$  telles que  $x_0 = a$ , montrer que la mesure spectrale de la fonction  $f = \mathbb{1}_{[a]} - \mu([a])$  (qui est la fonction indicatrice du cylindre  $[a]$  additivement renormalisée pour appartenir à  $\mathbb{L}_0^2(\mu)$ ) est un multiple de la mesure de Haar du cercle.

En étendant la transformation de Fourier des mesures positives sur le cercle aux mesures complexes, calculer la transformation de Fourier de la mesure complexe  $z^n d\mu_{\mathbb{S}_1}(z)$  pour tout  $n \in \mathbb{Z}$ . Montrer que pour tout cylindre non vide  $C$ , la mesure spectrale de  $f = \mathbb{1}_C - \mu(C)$  est dans la même classe de mesure<sup>99</sup> que la mesure de Haar  $\mu_{\mathbb{S}_1}$  du cercle.

99. c'est-à-dire qu'elles sont absolument continues l'une par rapport à l'autre

**Exercice E.21.** Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret inversible. Montrer que les assertions suivantes sont équivalentes.

- (1) Le système dynamique  $(X, \mathcal{B}, \mu, \phi)$  est *faiblement mélangeant*, c'est-à-dire que pour tous les  $A, B \in \mathcal{B}$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(\phi^{-k}(A) \cap B) - \mu(A)\mu(B)| = 0.$$

- (2) Pour tout  $f \in \mathbb{L}_0^2(\mu)$ , la mesure spectrale  $\sigma_f$  n'a pas d'atome.  
 (3) Le spectre ponctuel  $\text{Vp}(U_\phi|_{\mathbb{L}_0^2(\mu)})$  de la restriction de  $U_\phi$  à  $\mathbb{L}_0^2(\mu)$  est vide.  
 (4) Les seules fonctions propres de  $U_\phi$  sont les fonctions (presque partout) constantes.

En déduire qu'une transformation mélangeante est faiblement mélangeante et qu'une rotation irrationnelle du cercle n'est pas faiblement mélangeante.

### 3.4 Exercices

**Exercice E.22.** Soient  $(X, \mathcal{B}, \mu, \phi)$  et  $(X', \mathcal{B}', \mu', \phi')$  deux systèmes dynamiques mesurés, avec  $\mu$  et  $\mu'$  des mesures de probabilité. Soit  $h : X \rightarrow X'$  une semi-conjugaison entre  $(X, \phi)$  et  $(X', \phi')$ . Parmi les propriétés suivantes, quelles sont celles qui sont vérifiées par  $(X', \phi')$  si elles le sont par  $(X, \phi)$ , éventuellement quitte à rajouter des hypothèses sur  $h$  ?

- être ergodique,
- être mélangeant,
- être mélangeant d'ordre  $k$  pour un  $k \in \mathbb{N} - \{0, 1\}$  donné (voir l'exercice E.23),
- vérifier la propriété de mélange multiple (voir l'exercice E.23),
- être  $\psi$ -mélangeant sur un sous-espace vectoriel dense.

**Exercice E.23.** Soit  $(X, \mathcal{A}, \mu, \phi)$  un système dynamique probabilisé. Pour tout entier  $k \in \mathbb{N} \setminus \{0, 1\}$ , nous dirons que  $(X, \mathcal{A}, \mu, \phi)$  (ou  $\phi$  lorsque  $(X, \mathcal{A}, \mu)$  est sous-entendu ou  $\mu$  lorsque  $(X, \mathcal{A}, \phi)$  est sous-entendu) est *mélangeant d'ordre  $k$* <sup>100</sup> si pour toutes les parties mesurables  $A_1, \dots, A_k$  de  $X$ , et toutes les applications  $\tau_1, \dots, \tau_k$  de  $\mathbb{N}$  dans  $\mathbb{N}$  si le temps est discret, ou de  $\mathbb{R}$  dans  $\mathbb{R}$  si le temps est continu, telles que  $\lim_{t \rightarrow +\infty} \tau_{i+1}(t) - \tau_i(t) = +\infty$  pour tout  $i = 1, \dots, k-1$ , nous avons

$$\mu\left(\bigcap_{i=1}^k \phi^{-\tau_i(t)}(A_i)\right) \xrightarrow{t \rightarrow +\infty} \prod_{i=1}^k \mu(A_i).$$

Nous dirons que  $(X, \mathcal{A}, \mu, \phi)$  (ou  $\phi$  lorsque  $(X, \mathcal{A}, \mu)$  est sous-entendu ou tout simplement  $\mu$  lorsque  $(X, \mathcal{A}, \phi)$  est sous-entendu) vérifie la propriété de *mélange multiple* s'il est mélangeant d'ordre  $k$  pour tout  $k$  dans  $\mathbb{N} - \{0, 1\}$ .

- (1) Montrer que  $(X, \mathcal{A}, \mu, \phi)$  est mélangeant d'ordre  $k$  si et seulement si l'une des deux assertions suivantes est vérifiée

---

100. Remarquons que la propriété de mélange d'ordre 2 est équivalente à la propriété de mélange, par invariance de la mesure  $\mu$  par  $\phi$ . De plus, en prenant pour  $f_1$  la fonction constante 1, si  $(X, \mathcal{A}, \mu, \phi)$  est mélangeant d'ordre  $k \geq 3$ , alors il est mélangeant d'ordre  $k-1$ .

• pour tous les  $f_1, \dots, f_k \in \mathbb{L}^k(X, \mu)$  et toutes les applications  $\tau_1, \dots, \tau_k$  comme ci-dessus, nous avons

$$\int_X \left( \prod_{i=1}^k f_i \circ \phi^{\tau_i(t)} \right) d\mu \xrightarrow{t \rightarrow +\infty} \prod_{i=1}^k \int_X f_i d\mu. \quad (32)$$

• il existe une partie  $P$  de  $\mathbb{L}^k(X, \mu)$ , qui engendre un sous-espace vectoriel dense  $E$  de  $\mathbb{L}^k(X, \mu)$ , telle que pour tous les  $f_1, \dots, f_k$  dans  $P$  et toutes les applications  $\tau_1, \dots, \tau_k$  comme ci-dessus, la formule (32) soit vérifiée.

(2) Soient  $\mathcal{A}$  un alphabet dénombrable (muni de la topologie discrète) et  $\mu$  une mesure de probabilité sur  $\mathcal{A}$ . Montrer que le système de Bernoulli sur  $\mathcal{A}$ , muni de la mesure produit  $\mu^{\mathbb{N}}$  dans le cas unilatère et de la mesure produit  $\mu^{\mathbb{Z}}$  dans le cas bilatère, vérifie la propriété de mélange multiple.

(3) Soit  $p$  dans  $\mathbb{Z} - \{-1, 0, 1\}$ . Montrer que l'application  $\phi_p : x \mapsto p x$  du cercle  $\mathbb{R}/\mathbb{Z}$  dans lui-même vérifie la propriété de mélange multiple pour la mesure de Lebesgue du cercle.

(4) Soient  $N$  dans  $\mathbb{N} - \{0\}$  et  $M \in \mathcal{M}_N(\mathbb{Z})$  une matrice  $N \times N$  à coefficients entiers de déterminant non nul, dont les valeurs propres (complexes) sont de module strictement supérieur à 1. Montrer que l'application  $\phi_M : x \bmod \mathbb{Z}^N \mapsto Mx \bmod \mathbb{Z}^N$  du tore  $\mathbb{T}^N = \mathbb{R}^N / \mathbb{Z}^N$  dans lui-même vérifie la propriété de mélange multiple pour la mesure de Haar du tore.

**Exercice E.24.** Soient  $(X, \mu, \phi)$  et  $(X', \mu', \phi')$  deux systèmes dynamiques probabilisés, tous deux à temps continu ou à temps discret. Nous noterons  $\phi \times \phi'$  l'application ou le flot de l'espace mesurable produit  $X \times X'$  dans lui-même défini terme par terme par  $(x, x') \mapsto (\phi(x), \phi'(x'))$ .

- (1) Montrer que si  $(X, \mu, \phi)$  et  $(X', \mu', \phi')$  sont mélangeants, alors le système dynamique mesuré produit  $(X \times X', \mu \otimes \mu', \phi \times \phi')$  est mélangeant.
- (2) Soit  $k \in \mathbb{N} - \{0\}$ . Montrer que si  $(X, \mu, \phi)$  et  $(X', \mu', \phi')$  sont mélangeants d'ordre  $k$ , alors le système dynamique mesuré produit  $(X \times X', \mu \otimes \mu', \phi \times \phi')$  est mélangeant d'ordre  $k$ .
- (3) Soit  $E$  (respectivement  $E'$ ) un sous-espace vectoriel dense de  $\mathbb{L}^2(X, \mu)$  (respectivement  $\mathbb{L}^2(X', \mu')$ ), muni d'une norme  $\| \cdot \|_E$  (respectivement  $\| \cdot \|_{E'}$ ) telle que  $\| \cdot \|_{\mathbb{L}^2} \leq c \| \cdot \|_E$  (respectivement  $\| \cdot \|_{\mathbb{L}^2} \leq c \| \cdot \|_{E'}$ ) où  $c > 0$  est une constante. Soit  $\psi : \mathbb{N} \rightarrow ]0, 1]$  si  $(X, \mathcal{B}, \mu, \phi)$  est à temps discret, et  $\psi : \mathbb{R} \rightarrow ]0, 1]$  si  $(X, \mathcal{B}, \mu, \phi)$  est à temps continu, une application telle que  $\lim_{t \rightarrow +\infty} \psi(t) = 0$ . Soit  $E''$  le sous-espace vectoriel de  $\mathbb{L}^2(X \times X', \mu \otimes \mu')$  engendré par les applications à variables séparées définies par  $f \cdot f' : (x, x') \mapsto f(x) f'(x')$ , où  $f \in E$  et  $f' \in E'$ . Munissons  $E''$  d'une norme équivalente à la norme

$$\|g\|_{E''} = \inf \left\{ \sum_{i=1}^n \|f_i\|_E \|f'_i\|_{E'} : f_i \in E, f'_i \in E' \text{ et } g = \sum_{i=1}^n f_i \cdot f'_i \right\}.$$

Montrer que si  $(X, \mu, \phi)$  est  $\psi$ -mélangeant sur  $E$  et si  $(X', \mu', \phi')$  est  $\psi$ -mélangeant sur  $E'$ , alors  $(X \times X', \mu \otimes \mu', \phi \times \phi')$  est  $\psi$ -mélangeant sur  $E''$ .

### 3.5 Indications pour la résolution des exercices

**Correction de l'exercice E.19.** Les éléments de  $\ker(U_\phi - \text{id})$  sont les fonctions  $\mathbb{L}^2$  et presque partout invariants, donc les fonctions  $\mathbb{L}^2$  et  $\mathcal{I}$ -mesurables.

**Correction de l'exercice E.20.** Soient  $C$  un cylindre,  $k \in \mathbb{N} \setminus \{0\}$  sa longueur et  $c = \mu(C)$  sa mesure. Pour tout  $m \in \mathbb{Z}$ , si  $f = \mathbb{1}_C - c$  nous avons

$$\begin{aligned} c_m(\sigma_f) &= \langle f \circ \phi^m, f \rangle = \langle \mathbb{1}_C \circ \phi^m, \mathbb{1}_C \rangle - \langle c, \mathbb{1}_C \rangle - \langle \mathbb{1}_C \circ \phi^m, c \rangle + \langle c, c \rangle \\ &= \mu(\phi^{-m}(C) \cap C) - 2c\mu(C) + c^2 = \mu(\phi^{-m}(C) \cap C) - c^2. \end{aligned}$$

Rappelons le calcul de la transformation de Fourier de la mesure de Haar du cercle donné dans la formule (30). Par le calcul précédent, nous avons donc

$$c_0(\sigma_f) = c(1 - c) = c_0(c(1 - c) \mu_{\mathbb{S}_1})$$

et si  $|m| \geq k$ , alors  $c_m(\sigma_f) = 0 = c_m(\mu_{\mathbb{S}_1})$  car le cylindre  $C$  et le cylindre décalé  $\phi^{-m}(C)$  à gauche  $m$  fois si  $m > 0$  et à droite  $-m$  fois si  $m < 0$  sont alors de support disjoint (voir la démonstration de la proposition 3.2). Donc si  $k = 1$ , alors par l'injectivité de la transformation de Fourier des mesures positives sur  $\mathbb{S}_1$ , nous avons

$$\sigma_f = c(1 - c) \mu_{\mathbb{S}_1}.$$

La transformation de Fourier  $\mu \mapsto (c_m(\mu))_{m \in \mathbb{Z}}$  est encore définie sur l'espace vectoriel complexe  $\mathcal{M}_{\mathbb{C}}(\mathbb{S}_1)$  des mesures complexes sur  $\mathbb{S}_1$ , muni de la topologie faible-étoile, à valeur dans l'espace vectoriel topologique produit  $\mathbb{C}^{\mathbb{Z}}$ . Elle est encore continue, injective, et un homéomorphisme sur son image. De plus, la transformation de Fourier est linéaire

$$\forall m \in \mathbb{Z}, \forall \lambda \in \mathbb{C}, \forall \mu, \nu \in \mathcal{M}_{\mathbb{C}}(\mathbb{S}_1) \quad c_m(\lambda\mu + \nu) = \lambda c_m(\mu) + c_m(\nu).$$

Pour tous les  $m, n \in \mathbb{Z}$  et  $\mu \in \mathcal{M}_{\mathbb{C}}(\mathbb{S}_1)$ , nous avons

$$c_m(z^n \mu) = \int_{z \in \mathbb{S}_1} z^{-m} z^n d\mu(z) = c_{m-n}(\mu).$$

En particulier, nous avons  $c_m(z^n \mu_{\mathbb{S}_1}) = \delta_{n,m}$ . Pour tout  $n \in \llbracket -k+1, k-1 \rrbracket$ , posons  $a_n = \mu(\phi^{-n}(C) \cap C) - c^2$ , en remarquant que  $a_{-n} = a_n$  par l'invariance de la mesure. Notons  $\rho : \mathbb{S}_1 \rightarrow \mathbb{C}$  le polynôme trigonométrique  $z \mapsto \sum_{n=-k+1}^{k-1} a_n z^n$ , en remarquant qu'il n'admet qu'un nombre fini de 0 dans  $\mathbb{S}_1$ . Alors par la linéarité de la transformation de Fourier des mesures complexes sur le cercle, pour tout  $m, n \in \mathbb{Z}$ , nous avons

$$c_m(\rho \mu_{\mathbb{S}_1}) = \sum_{n=-k+1}^{k-1} a_n c_m(z^n \mu_{\mathbb{S}_1}) = \sum_{n=-k+1}^{k-1} a_n \delta_{n,m} = \begin{cases} 0 = c_m(\sigma_f) & \text{si } |m| \geq k \\ a_m = c_m(\sigma_f) & \text{sinon.} \end{cases}$$

Par l'injectivité de la transformation de Fourier des mesures complexes sur le cercle, nous avons donc  $\sigma_f = \rho \mu_{\mathbb{S}_1}$ . Puisque  $\rho$  n'a qu'un nombre fini de zéros sur le cercle, les mesures  $\sigma_f = \rho \mu_{\mathbb{S}_1}$  sont bien absolument continues l'une par rapport à l'autre. <sup>101</sup>

101. Puisque  $\sigma_f$  et  $\mu_{\mathbb{S}_1}$  sont deux mesures positives, nous avons  $\rho(z) \geq 0$  pour tout  $z \in \mathbb{S}_1$ , ce qui peut se montrer directement en utilisant le fait que  $z^n + z^{-n} + 2 = 2\text{Re } z^n + 2 \geq 0$ . Il est aussi possible de montrer ce résultat en n'utilisant que la transformation de Fourier des mesures positives sur le cercle, mais la démonstration est moins naturelle.

Par exemple, si  $C = [a, b]$  est un cylindre de longueur  $k = 2$  avec  $a, b$  distincts dans  $\mathcal{A}$ , alors  $\phi^{\pm 1}(C) \cap C$  est vide, donc  $a_1 = a_{-1} = -c^2$  et  $\rho(e^{2i\pi\theta}) = a_{-1}e^{-2i\pi\theta} + a_0 + a_1e^{2i\pi\theta} = c - c^2(1 + 2\cos(2\pi\theta))$ .

**Correction de l'exercice E.21 .** Les assertions (3) et (4) sont clairement équivalentes. Le fait qu'un système dynamique mélangeant est faiblement mélangeant a été vu après l'énoncé de la proposition 3.4. Le fait qu'une rotation irrationnelle ne soit pas faiblement mélangeante découle du calcul du spectre de son opérateur de Koopman, voir l'exemple (2) du début de la partie 3.3.

**Correction de l'exercice E.22 .** • Supposons que  $(X, \mathcal{B}, \mu, \phi)$  soit ergodique. Si une partie mesurable  $A' \in \mathcal{B}'$  est invariante par  $\phi'$ , alors

$$\phi^{-1}(h^{-1}(A')) = (h \circ \phi)^{-1}(A') = (\phi' \circ h)^{-1}(A') = h^{-1}(\phi'^{-1}(A')) = h^{-1}(A') .$$

Puisque  $h$  est mesurable, la partie mesurable  $h^{-1}(A') \in \mathcal{B}$  est donc invariante par  $\phi$ . Puisque  $h$  préserve les mesures (c'est-à-dire  $\mu' = h_*\mu$ ) et par ergodicité, nous avons donc

$$\mu'(A') = h_*\mu(A') = \mu(h^{-1}(A')) \in \{0, 1\} .$$

Donc  $(X', \mathcal{B}', \mu', \phi')$  est ergodique.

• Rappelons qu'un système dynamique mesuré est mélangeant d'ordre 2 si et seulement s'il est mélangeant, et que par définition il vérifie la propriété de mélange multiple s'il est mélangeant d'ordre  $k$  pour  $k$  dans  $\mathbb{N} - \{0, 1\}$ . Il suffit donc de montrer le troisième point pour obtenir le deuxième et le quatrième. Nous supposons que le temps est discret, la démonstration dans le cas du temps continu est similaire.

Soit  $k \in \mathbb{N} - \{0, 1\}$ , supposons que  $(X, \mathcal{B}, \mu, \phi)$  soit mélangeant d'ordre  $k$ . Soient  $A'_1, \dots, A'_k$  des parties mesurables de  $X'$  et  $\tau_1, \dots, \tau_k$  des applications  $\mathbb{N}$  dans  $\mathbb{N}$  telles que  $\lim_{n \rightarrow +\infty} \tau_{i+1}(n) - \tau_i(n) = +\infty$  pour  $i = 1, \dots, k-1$ . Alors  $h^{-1}(A'_1), \dots, h^{-1}(A'_k)$  sont des parties mesurables de  $X$ , par la mesurabilité de  $h$ . Puisque  $(X', \mathcal{B}', \mu', \phi')$  est mélangeant d'ordre  $k$ , et puisque  $h$  préserve les mesures (c'est-à-dire  $\mu' = h_*\mu$ ), nous avons

$$\begin{aligned} \mu' \left( \bigcap_{i=1}^k \phi'^{-\tau_i(n)}(A'_i) \right) &= \mu \left( h^{-1} \left( \bigcap_{i=1}^k \phi'^{-\tau_i(n)}(A'_i) \right) \right) = \mu \left( \bigcap_{i=1}^k \phi^{-\tau_i(n)}(h^{-1}(A'_i)) \right) \\ &\xrightarrow{n \rightarrow +\infty} \prod_{i=1}^k \mu(h^{-1}(A'_i)) = \prod_{i=1}^k \mu'(A'_i) . \end{aligned}$$

Donc  $(X', \mathcal{B}', \mu', \phi')$  est mélangeant d'ordre  $k$ .

• Montrons le dernier point. Nous supposons que le temps est discret, la démonstration dans le cas du temps continu est similaire. Soit  $\psi : \mathbb{N} \rightarrow ]0, +\infty[$  une application telle que  $\lim_{n \rightarrow +\infty} \psi(n) = 0$ . Soient  $E$  et  $E'$  deux sous-espaces vectoriels denses de  $\mathbb{L}^2(X, \mu)$  et  $\mathbb{L}^2(X', \mu')$ , munis de normes  $\| \cdot \|_E$  et  $\| \cdot \|_{E'}$ , respectivement. Faisons l'hypothèse supplémentaire que l'application linéaire  $L_h : f \mapsto f \circ h$  de précomposition par  $h$  envoie  $E'$  dans  $E$ , et qu'elle est continue pour les normes de  $E'$  et de  $E$ .<sup>102</sup> Soit donc  $\kappa > 0$  tel que  $\|f \circ h\|_E \leq \kappa \|f\|_{E'}$  pour tout  $f \in E'$ .

102. Par exemple, si  $E = \mathbb{L}^2(X, \mu)$  et  $E' = \mathbb{L}^2(X', \mu')$ , alors le fait que  $h$  préserve les mesures montre (et en fait est équivalent au fait) que l'application  $L_h$  de précomposition par  $h$  préserve la norme : nous avons  $\|f \circ h\|_E = \|f\|_{E'}$  pour tout  $f \in E'$ .

Supposons que  $(X, \mathcal{B}, \mu, \phi)$  soit  $\psi$ -mélangeant sur  $E$ . Pour tous les  $f', g' \in E'$ , et pour tout  $n \in \mathbb{N}$ , avec la constante  $c$  donnée par la propriété de  $\psi$ -mélange de  $(X, \mathcal{B}, \mu, \phi)$ , puisque  $h$  préserve les mesures, nous avons

$$\begin{aligned} & \left| \int_{X'} (f' \circ \phi'^n) g' d\mu' - \left( \int_{X'} f' d\mu' \right) \left( \int_{X'} g' d\mu' \right) \right| \\ &= \left| \int_X (f' \circ \phi'^n \circ h) (g' \circ h) d\mu - \left( \int_X f' \circ h d\mu \right) \left( \int_X g' \circ h d\mu \right) \right| \\ &= \left| \int_X ((f' \circ h) \circ \phi^n) (g' \circ h) d\mu - \left( \int_X f' \circ h d\mu \right) \left( \int_X g' \circ h d\mu \right) \right| \\ &\leq c \|f' \circ h\|_E \|g' \circ h\|_E \psi(n) \leq c \kappa^2 \|f'\|_{E'} \|g'\|_{E'} \psi(n). \end{aligned}$$

Donc  $(X', \mathcal{B}', \mu', \phi')$  est  $\psi$ -mélangeant sur  $E'$ .

**Correction de l'exercice E.23.** (1) La démonstration est très proche de celle de la proposition 3.1. Par invariance de la mesure, si une fonction est de classe  $\mathbb{L}^k$ , alors sa composée par une puissance quelconque de  $\phi$  l'est encore. Rappelons que le produit de  $k$  fonctions de classe  $\mathbb{L}^k$  est de classe  $\mathbb{L}^1$  :

$$\left| \int_X f_1 \dots f_k d\mu \right| \leq \prod_{i=1}^k \|f_i\|_{\mathbb{L}^k} = \prod_{i=1}^k \left( \int_X |f_i|^k d\mu \right)^{1/k}. \quad (33)$$

Ceci se démontre par récurrence sur  $k$  en utilisant l'inégalité de Hölder

$$\left| \int_X f g d\mu \right| \leq \left( \int_X |f|^p d\mu \right)^{1/p} \left( \int_X |g|^q d\mu \right)^{1/q}$$

si  $\frac{1}{p} + \frac{1}{q} = 1$  si  $f$  est  $\mathbb{L}^p$  et  $g$  est  $\mathbb{L}^q$ , en prenant  $f = f_1 \dots f_{k-1}$ ,  $g = f_k$ ,  $q = k$  et  $p = \frac{k}{k-1}$ . Puisque la mesure  $\mu$  est une mesure de probabilité, une fonction  $h$  de classe  $\mathbb{L}^k$  est de classe  $\mathbb{L}^1$ , et  $\left| \int_X h d\mu \right| \leq \|h\|_{\mathbb{L}^k}$ . Donc les différents termes sont bien définis.

Le premier point implique trivialement le second, et il implique le mélange d'ordre  $k$ , car il suffit de prendre les fonctions indicatrices  $f_i = \mathbb{1}_{A_i}$  pour  $1 \leq i \leq k$ . Pour tout temps  $t$ , posons

$$c_t(f_1, \dots, f_k) = \int_X \left( \prod_{i=1}^k f_i \circ \phi^{\tau_i(t)} \right) d\mu - \prod_{i=1}^k \int_X f_i d\mu,$$

qui est  $k$ -linéaire en  $(f_1, \dots, f_k)$ . Si le système  $(X, \mathcal{B}, \mu, \phi)$  est mélangeant d'ordre  $k$ , alors par multilinéarité, la propriété  $\lim_{t \rightarrow \infty} c_t(f_1, \dots, f_k) = 0$  est vérifiée pour toutes les fonctions étagées  $f_1, \dots, f_k$ . Montrons que pour tous les  $f_1, \dots, f_{k-1}$  dans  $\mathbb{L}^k(X, \mu)$ , le sous-espace vectoriel  $H_{f_1, \dots, f_{k-1}}$  des fonctions  $f_k$  dans  $\mathbb{L}^k(X, \mu)$  telles que

$$\lim_{t \rightarrow \infty} c_t(f_1, \dots, f_k) = 0$$

est fermé dans  $\mathbb{L}^k(X, \mu)$ . Puisque l'espace vectoriel des fonctions étagées est dense dans  $\mathbb{L}^k(X, \mu)$ , et par une récurrence, cela montrera le résultat.

Nous pouvons supposer que  $f_1, \dots, f_{k-1}$  ne sont pas presque partout nulles, sinon le résultat est immédiat. Soit  $f_k \in \overline{H_{f_1, \dots, f_{k-1}}}$  et, pour tout  $\epsilon > 0$ , soit  $g_k \in H_{f_1, \dots, f_{k-1}}$  telle que  $\|f_k - g_k\|_{\mathbb{L}^k} \leq \frac{\epsilon}{3 \prod_{i=1}^{k-1} \|f_i\|_{\mathbb{L}^k}}$ . Pour tout temps  $t$  assez grand, puisque  $g_k \in H_{f_1, \dots, f_{k-1}}$ ,



nous avons  $|c_t(f_1, \dots, f_{k-1}, g_k)| \leq \frac{\epsilon}{3}$ . Par l'inégalité (33) et la majoration de la norme  $\mathbb{L}^1$  par la norme  $\mathbb{L}^k$ , nous avons

$$\begin{aligned} |c_t(f_1, \dots, f_{k-1}, f_k)| &= \left| \int_X \left( \prod_{i=1}^{k-1} f_i \circ \phi^{\tau_i(t)} \right) ((f_k - g_k) \circ \phi^{\tau_k(t)}) d\mu \right. \\ &\quad \left. - \left( \prod_{i=1}^{k-1} \int_X f_i d\mu \right) \left( \int_X (f_k - g_k) d\mu \right) + c_t(f_1, \dots, f_{k-1}, g_k) \right| \\ &\leq 2 \prod_{i=1}^{k-1} \|f_i\|_{\mathbb{L}^k} \|f_k - g_k\|_{\mathbb{L}^k} + |c_t(f_1, \dots, f_{k-1}, g_k)| \leq 2\frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

Donc  $f_k$  appartient à  $H_{f_1, \dots, f_{k-1}}$ , et  $H_{f_1, \dots, f_{k-1}}$  est fermé.

(2) Nous considérons le cas du décalage unilatère, le cas bilatère se traite de manière similaire. Soient  $\mathcal{A}$  un alphabet dénombrable discret,  $X = \Sigma_+ = \mathcal{A}^{\mathbb{N}}$  et  $\phi = \sigma_+$  le décalage à gauche sur  $X$ . Comme la tribu ( $\sigma$ -algèbre) de l'espace mesurable produit  $X$  est engendrée par les cylindres, il suffit de vérifier l'assertion lorsque les  $A_i$  sont des cylindres. Soient  $p_1, \dots, p_k$  dans  $\mathbb{N}$  et, pour tous les  $i = 1, \dots, k$  et  $j = 0, \dots, p_i$ , soit  $x_{j,i} \in \mathcal{A}$ . Notons  $A_i$  le cylindre  $[x_{0,i}, x_{1,i}, \dots, x_{p_i,i}]$ . Alors si  $n$  est assez grand, par l'hypothèse sur les fonctions  $\tau_i$ , les parties de  $\mathbb{N}$  définies par  $\{0, \dots, p_i\} + \tau_i(n)$  pour  $i = 1, \dots, k$  sont deux à deux disjointes. Donc, si  $n$  est assez grand, par définition du décalage et de la valeur de la mesure produit sur les cylindres,

$$\mu^{\mathbb{N}} \left( \bigcap_{i=1}^k \sigma_+^{-\tau_i(n)}(A_i) \right) = \prod_{i=1}^k \prod_{j=1}^{p_i} \mu(\{x_{j,i}\}) = \prod_{i=1}^k \mu^{\mathbb{N}}(A_i),$$

ce qui montre le résultat.

(3) et (4). L'assertion (3) découle de l'assertion (2) et du dernier point de l'exercice E.22 si  $p \geq 0$ , car l'application de développement en base  $p$  est une semi-conjugaison entre le système de Bernoulli unilatère  $(\mathcal{A}^{\mathbb{N}}, \mu^{\mathbb{N}}, \sigma_+)$  sur l'alphabet  $\{0, \dots, p-1\}$  muni de l'équiprobabilité  $\mu = \frac{1}{p} \sum_{i=0}^{p-1} \Delta_i$  (où  $\Delta_i$  est la masse de Dirac unité en  $i$ ) et le système dynamique mesuré  $(\mathbb{R}/\mathbb{Z}, d\theta, \phi_p)$ . Mais elle est aussi un cas particulier de l'assertion (4) avec  $N = 1$ .

Soient  $X = \mathbb{R}^N / \mathbb{Z}^N$  et  $\mu = \lambda = d\theta$  la mesure de Haar sur  $X$ . Pour tout  $m$  dans  $\mathbb{Z}^N$ , notons comme d'habitude  $e_m : X \rightarrow \mathbb{C}$  la fonction trigonométrique  $\theta \mapsto e^{2i\pi\langle m, \theta \rangle}$ . Par l'assertion (1) et par le théorème de Stone-Weierstrass, il suffit de montrer que pour tous les  $k \in \mathbb{N} - \{0\}$  et  $m_1, \dots, m_k$  dans  $\mathbb{Z}^N$ , la formule (32) est vérifiée pour  $\mu = \lambda$ ,  $f_j = e_{m_j}$  pour  $j = 1, \dots, k$  et  $\phi = \phi_M$ . Raisonnons par récurrence sur  $k$ . C'est vrai pour  $k = 1$ , comme conséquence de l'invariance de  $\lambda$  par  $\phi_M$ . Supposons  $k \geq 2$ . Si  $m_k = 0$ , alors le résultat est vrai par l'hypothèse au rang  $k-1$ . Supposons donc  $m_k$  non nul. En particulier,

$$\int_X f_k d\mu = 0 = \prod_{j=1}^k \int_X f_j d\lambda.$$

Nous avons  $f_j \circ \phi_M^{\tau_j(n)} = e_{m_{j,n}}$  avec  $m_{j,n} = {}^t M^{\tau_j(n)} m_j$ . Montrons que pour tout  $n$  assez

grand, la somme  $\sum_{j=1}^k {}^t M^{\tau_j(n)} m_j$  est non nulle. Ceci implique que

$$\int_X \prod_{j=1}^k f_j \circ \phi_M^{\tau_j(n)} d\lambda = \int_X e_{\sum_{j=1}^k {}^t M^{\tau_j(n)} m_j}(\theta) d\theta = 0,$$

ce qu'il fallait démontrer.

Supposons par l'absurde que

$$\sum_{j=1}^k {}^t M^{\tau_j(n)} m_j = 0 \tag{34}$$

pour une sous-suite d'entiers  $n$  tendant vers  $+\infty$ . En trigonalisant  $M$  sur  $\mathbb{C}$  et puisque les valeurs propres complexes de  $M$  sont de module strictement supérieur à 1, il existe une matrice triangulaire inférieure complexe  $T$ , dont nous pouvons supposer que les coefficients diagonaux  $\lambda_1, \dots, \lambda_N$  soient ordonnés de sorte que  $1 < |\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_N|$  et une matrice inversible  $P \in \text{GL}_n(\mathbb{C})$  telle que  $M = P T P^{-1}$ . Notons  $v_j = {}^t P m_j$  pour  $j = 1, \dots, k$ . La formule (34) implique alors

$$\sum_{j=1}^k {}^t T^{\tau_j(n)} v_j = 0 \tag{35}$$

pour une sous-suite d'entiers  $n$  tendant vers  $+\infty$ . Puisque le vecteur  $v_k = (v_{k,1}, \dots, v_{k,N})$  est non nul, il existe  $r \in \{1, \dots, N\}$  tel que  $v_{k,r} \neq 0$  et  $v_{k,\ell} = 0$  pour  $\ell = r+1, \dots, N$ . Puisque  ${}^t T$  est triangulaire supérieure, le module du  $r$ -ème coefficient du  $N$ -uplet  ${}^t T^{\tau_k(n)} v_k$  est égal à  $|\lambda_r|^{\tau_k(n)} |v_{k,r}|$ . Puisque  $\lim_{n \rightarrow +\infty} (\tau_{k-1}(n) - \tau_j(n)) = +\infty$  pour  $j = 1, \dots, k-2$ , le module du  $r$ -ème coefficient de  $\sum_{j=1}^{k-1} {}^t T^{\tau_j(n)} v_j$  est borné par  $c |\lambda_r|^{\tau_{k-1}(n)}$  pour une constante  $c \geq 0$ . Mais comme  $\lim_{n \rightarrow +\infty} (\tau_{k-1}(n) - \tau_{k-1}(n)) = +\infty$  et  $|\lambda_r| > 1$ , ceci contredit la formule (35).

**Correction de l'exercice E.24.** Nous supposons que le temps est discret, le temps continu se traite de manière semblable.

(1) C'est le cas particulier  $k = 2$  de l'assertion suivante.

(2) Soient  $\tau_1, \dots, \tau_k : \mathbb{N} \rightarrow \mathbb{N}$  telles que  $\lim_{n \rightarrow +\infty} (\tau_{i+1} - \tau_i) = +\infty$  pour  $i = 1, \dots, k-1$ .

Montrons que pour tous les boréliens  $A_1'', \dots, A_k''$  de  $X \times X'$ , nous avons

$$\lim_{n \rightarrow +\infty} \mu \otimes \mu' \left( \bigcap_{i=1}^k (\phi \times \phi')^{-\tau_i(n)} (A_i'') \right) = \prod_{i=1}^k \mu \otimes \mu' (A_i'').$$

Puisque les produits de boréliens engendrent la tribu borélienne produit, nous pouvons supposer que  $A_i'' = A_i \times A_i'$  où  $A_i$  est un borélien de  $X$  et  $A_i'$  un borélien de  $X'$ , pour tout

$i = 1, \dots, k$ . Nous avons alors

$$\begin{aligned}
\mu \otimes \mu' \left( \bigcap_{i=1}^k (\phi \times \phi')^{-\tau_i(n)}(A_i'') \right) &= \mu \otimes \mu' \left( \bigcap_{i=1}^k \phi^{-\tau_i(n)}(A_i) \times \bigcap_{i=1}^k \phi'^{-\tau_i(n)}(A_i') \right) \\
&= \mu \left( \bigcap_{i=1}^k \phi^{-\tau_i(n)}(A_i) \right) \mu' \left( \bigcap_{i=1}^k \phi'^{-\tau_i(n)}(A_i') \right) \\
&\xrightarrow{n \rightarrow +\infty} \prod_{i=1}^k \mu(A_i) \prod_{i=1}^k \mu'(A_i') = \prod_{i=1}^k \mu \otimes \mu'(A_i'').
\end{aligned}$$

(3) Notons que

$$E'' = \left\{ g : X \times X' \rightarrow \mathbb{C} : \exists n \in \mathbb{N}, f_i \in E, f_i' \in E', g = \sum_{i=1}^n f_i \cdot f_i' \right\}$$

est bien un sous-espace vectoriel de  $\mathbb{L}^2(X \times X', \mu \otimes \mu')$ , car stable par addition par définition, stable par multiplication scalaire car  $E$  l'est, et composé de fonctions de classe  $\mathbb{L}^2$  par la majoration  $\|g\|_{\mathbb{L}^2} \leq 2 \sum_{i=1}^n \|f_i\|_{\mathbb{L}^2} \|f_i'\|_{\mathbb{L}^2}$ . Il est de plus dense dans  $\mathbb{L}^2(X \times X', \mu \otimes \mu')$ , car le sous-espace vectoriel engendré par les fonctions à variables séparées est dense dans  $\mathbb{L}^2(X \times X', \mu \otimes \mu')$ , et que  $E$  et  $E'$  sont denses dans  $\mathbb{L}^2(X, \mu)$  et  $\mathbb{L}^2(X, \mu')$  respectivement.

En rappelant la notation  $c_{n,\phi}(f, g)$  pour le coefficient de corrélation

$$c_{n,\phi}(f, g) = \int_X (f \circ \phi^n) g \, d\mu - \left( \int_X f \, d\mu \right) \left( \int_X g \, d\mu \right),$$

montrons que pour tout  $n \in \mathbb{N}$

$$|c_{n,\phi \times \phi'}(f \cdot f', g \cdot g')| = O(\psi(n) \|f\|_E \|f'\|_{E'} \|g\|_E \|g'\|_{E'}).$$

Par la bilinéarité des coefficients de corrélation, et par la définition de la norme de  $E''$ , ceci montrera le résultat.

Puisque  $(f \cdot f') \circ (\phi \times \phi')^n = (f \circ \phi^n) \cdot (f' \circ \phi'^n)$  et  $\mu \otimes \mu'(h \cdot h') = \mu(h) \mu'(h')$  pour tous les  $h \in \mathbb{L}^2(X, \mu)$  et  $h' \in \mathbb{L}^2(X', \mu')$ , nous avons

$$\begin{aligned}
\int_{X \times X'} (f \cdot f') \circ (\phi \times \phi')^n (g \cdot g') \, d\mu \, d\mu' &= \left( \int_X (f \circ \phi^n) g \, d\mu \right) \left( \int_{X'} (f' \circ \phi'^n) g' \, d\mu' \right) \\
&= (\mu(f) \mu(g) + O(\psi(n) \|f\|_E \|g\|_E)) (\mu'(f') \mu'(g') + O(\psi(n) \|f'\|_{E'} \|g'\|_{E'}))
\end{aligned}$$

Pour tout élément  $f \in E$ , puisque  $\mu$  est une mesure de probabilité sur  $X$ , nous avons  $|\mu(f)| \leq \|f\|_{\mathbb{L}^2} \leq c \|f\|_E$  par l'inégalité de Cauchy-Schwarz et par l'hypothèse sur la norme de  $E$ . De même, pour tout  $f' \in E'$ , nous avons la majoration  $|\mu'(f')| \leq c \|f'\|_{E'}$ . Puisque  $\psi(n)^2 \leq \psi(n)$  car  $\psi \leq 1$  par hypothèse, nous avons donc, comme voulu,

$$|c_{n,\phi \times \phi'}(f \cdot f', g \cdot g')| = O(\psi(n) \|f\|_E \|g\|_E \|f'\|_{E'} \|g'\|_{E'}).$$

## 4 Dynamique linéaire mesurée des tores

Fixons  $N \in \mathbb{N} \setminus \{0\}$ . Le *tore*  $\mathbb{T}^N$  est le groupe topologique quotient de  $\mathbb{R}^N$  par son sous-groupe (distingué fermé)  $\mathbb{Z}^N$ , ou de manière équivalente l'ensemble des  $N$ -uplets  $\theta = (\theta_1, \dots, \theta_N)$  où  $\theta_i$  est un réel modulo 1, muni de la topologie produit et de l'addition modulo 1 terme à terme. C'est un groupe topologique métrisable compact et commutatif.

Les *translations* de  $\mathbb{T}^N$  sont les applications  $\tau_\theta : x \mapsto x + \theta$  pour  $\theta \in \mathbb{T}^N$ . Pour tout  $x \in \mathbb{R}^N$ , nous noterons  $\dot{x}$  la classe de  $x$  modulo  $\mathbb{Z}^N$ , et  $\rho : \mathbb{R}^N \rightarrow \mathbb{T}^N$  la *projection canonique*  $x \mapsto \dot{x}$ .

Il existe sur  $\mathbb{T}^N$  une unique mesure de probabilité borélienne invariante par toutes les translations : la *mesure de Haar* (aussi appelée mesure de Lebesgue)  $\lambda = d\theta = \mu_{\mathbb{T}^N}$  de  $\mathbb{T}^N$ , image par  $\rho : x \mapsto \dot{x}$  de la restriction au cube unité de la mesure de Lebesgue  $dx$  de  $\mathbb{R}^N$ , définie par

$$\forall f \in C^0(\mathbb{T}^N; \mathbb{C}), \quad \int_{\mathbb{T}^N} f(\theta) d\theta = \int_{[0,1]^N} f(\dot{x}) dx.$$

L'invariance par translations de  $\lambda$  découle de l'invariance par translations de la mesure de Lebesgue de  $\mathbb{R}^N$  car  $\tau_y(\dot{x}) = \rho(x + y)$  pour tous les  $x, y$  dans  $\mathbb{R}^N$ . L'unicité de  $\lambda$  découle du fait que la mesure de Lebesgue de  $\mathbb{R}^n$  est l'unique mesure borélienne positive sur  $\mathbb{R}^n$  invariante par translations, qui donne mesure 1 au cube unité  $[0, 1]^N$ .

Pour tout  $k \in \mathbb{N}$ , une application continue  $f : \mathbb{T}^N \rightarrow \mathbb{R}$  est dite *différentiable de classe  $C^k$*  si l'application  $\mathbb{Z}^N$ -invariante  $f \circ \rho : \mathbb{R}^N \rightarrow \mathbb{R}$  l'est. Pour tout  $\dot{x} \in \mathbb{T}^N$ , et pour tous les  $\ell \in \{1, \dots, k\}$  et  $i_1, \dots, i_\ell \in \{1, \dots, N\}$ , les *dérivées partielles* de  $f$  d'ordre  $\ell$  sont les applications

$$\partial_{i_1, \dots, i_\ell} f : \dot{x} \mapsto \partial_{i_1, \dots, i_\ell} (f \circ \rho)(x) = \frac{\partial^\ell (f \circ \rho)}{\partial x_{i_1} \dots \partial x_{i_\ell}}(x),$$

ce qui ne dépend pas du choix du relevé  $x$  de  $\dot{x}$ .

Nous allons maintenant commencer à étudier des systèmes dynamiques mesurés sur le tore  $\mathbb{T}^N$  (voir la partie 11.4 pour la suite de l'étude). Cette étude sera facilitée par la possibilité d'utiliser de l'analyse de Fourier (voir par exemple [Die2, Chap XIX] pour toute information).

Nous renvoyons à la partie 8 pour une étude de dynamique topologique et différentiable dans le cas  $N = 1$ .

### 4.1 Séries de Fourier sur le tore

Pour  $m = (m_1, \dots, m_N)$  dans  $\mathbb{Z}^N$ , notons  $e_m \in C^0(\mathbb{T}^N; \mathbb{C})$  la *fonction trigonométrique* définie par

$$e_m : \theta \mapsto e^{2i\pi \langle m, \theta \rangle}$$

où  $\langle m, \theta \rangle = \sum_{j=1}^N m_j \theta_j \in \mathbb{R}/\mathbb{Z}$ . Remarquons que, pour tous les  $m, n \in \mathbb{Z}^N$ , nous avons

$$\overline{e_m} = e_{-m} \quad \text{et} \quad e_m e_n = e_{m+n}. \quad (36)$$

Pour toute mesure borélienne finie  $\mu$  sur  $\mathbb{T}^N$ , notons<sup>103</sup>  $(c_m(\mu))_{m \in \mathbb{Z}^N}$  la famille indexée par  $\mathbb{Z}^N$  de ses *coefficients de Fourier*, où

$$c_m(\mu) = \int_{\mathbb{T}^N} e_{-m} d\mu = \int_{\mathbb{T}^N} e^{-2i\pi \langle m, \theta \rangle} d\mu(\theta).$$

103. Voir la proposition 3.3 et les définitions qui la précède si  $N = 1$ .

Pour toute application continue  $f \in C^0(\mathbb{T}^N; \mathbb{C})$ , notons  $(c_m(f))_{m \in \mathbb{Z}^N}$  la famille indexée par  $\mathbb{Z}^N$  de ses *coefficients de Fourier*, où

$$c_m(f) = \int_{\mathbb{T}^N} e_{-m}(\theta) f(\theta) d\theta .$$

qui vérifient, pour tout  $m \in \mathbb{Z}^N$ , que  $c_m(f) = c_m(f d\theta)$  si  $f \geq 0$  (pour que  $f d\theta$  soit une mesure positive), et

$$|c_m(f)| \leq \|f\|_\infty .$$

**Proposition 4.1.** *Nous avons les propriétés suivantes.*

- (1) *Le sous-espace vectoriel complexe de  $C^0(\mathbb{T}^N; \mathbb{C})$  engendré par la famille trigonométrique  $(e_m)_{m \in \mathbb{Z}^N}$  est dense dans  $C^0(\mathbb{T}^N; \mathbb{C})$ .*
- (2) *La transformation de Fourier  $\mu \mapsto (c_m(\mu))_{m \in \mathbb{Z}^N}$  est injective.*
- (3) *Si  $f \in C^1(\mathbb{T}^N; \mathbb{C})$ , nous avons, avec convergence normale donc uniforme de la série,*

$$f(\theta) = \sum_{m \in \mathbb{Z}^N} c_m(f) e_m(\theta) .$$

- (4) *Si  $f \in C^1(\mathbb{T}^N; \mathbb{C})$ , pour tout  $m \in \mathbb{Z}^N$  et pour tout  $j = 1, \dots, N$ , en notant  $m(j) \in \mathbb{Z}$  la  $j$ -ème composante de  $m$  et en supposant qu'elle est non nulle, nous avons*

$$|c_m(f)| = \left| \frac{1}{2\pi m(j)} c_m(\partial_j f) \right| \leq \frac{1}{2\pi |m(j)|} \|\partial_j f\|_\infty .$$

- (5) *Pour tous les  $m, m' \in \mathbb{Z}^N$ , nous avons les formules d'orthogonalité*

$$\int_{\mathbb{T}^N} e_m \overline{e_{m'}} d\theta = \int_{\mathbb{T}^N} e_{m-m'} d\theta = \begin{cases} 1 & \text{si } m = m' \\ 0 & \text{sinon.} \end{cases} \quad (37)$$

**Démonstration.** (1) Le sous-espace vectoriel complexe de  $C^0(\mathbb{T}^N; \mathbb{C})$  engendré par la famille trigonométrique  $(e_m)_{m \in \mathbb{Z}^N}$  est une sous-algèbre, invariante par la conjugaison complexe, de  $C^0(\mathbb{T}^N; \mathbb{C})$  (voir la formule (36)). L'assertion (1) découle donc du théorème de Stone<sup>104</sup>-Weierstrass<sup>104, 105</sup>.

(2) Ceci découle du résultat ci-dessus de densité, et du théorème de représentation de Riesz.

(3) C'est une propriété de la transformation de Fourier inverse, voir [Die2, Chap XIX].

(4) Ceci découle d'une intégration par partie sur la  $j$ -ème composante.

(5) Le calcul est immédiat, par périodicité. □



Marshall Stone  
1903-1989



Karl Weierstrass  
1803-1896



Claude Shannon  
1916-2001



Yakov Sinai  
1935-



Hermann Weyl  
1885-1955

104. Le théorème de Stone-Weierstrass est le résultat de densité suivant (voir par exemple [Die1] ou [Pau3, §5.6]). Un ensemble  $\mathcal{A}$  d'applications d'un ensemble  $E$  dans  $\mathbb{C}$  est dit *séparant* si pour tous les  $x \neq y$  dans  $E$ , il existe  $f \in \mathcal{A}$  telle que  $f(x) \neq f(y)$ . L'espace vectoriel complexe  $C^0(X; \mathbb{C})$  est une algèbre pour la multiplication point par point des fonctions à valeurs complexes.

Soit  $X$  un espace topologique métrisable compact non vide. Toute sous-algèbre séparante et invariante par la conjugaison complexe de l'algèbre  $C^0(X; \mathbb{C})$  des applications continues de  $X$  dans  $\mathbb{C}$  est dense dans  $C^0(X; \mathbb{C})$  pour la topologie de la norme uniforme.

## 4.2 Unique ergodicité

Donnons deux motivations au développement de cette partie.

Une mesure de probabilité invariante, dont l'existence est assurée par exemple sous les hypothèses de la proposition 0.6, est rarement unique (considérer par exemple le cas où la transformation  $\phi$  vaut l'identité). Nous donnons ci-dessous un critère d'unicité.

Rappelons que le théorème de Birkhoff ne donne une convergence des moyennes de Birkhoff que presque partout pour une mesure invariante fixée. Cette convergence est rarement vérifiée en tout point. Nous donnons ci-dessous un critère de convergence ponctuelle des moyennes de Birkhoff de fonctions continues en vraiment tout point.

**Proposition 4.2.** *Soient  $X$  un espace topologique compact métrisable non vide, muni d'une application continue  $\phi : X \rightarrow X$ . Les assertions suivantes sont équivalentes.*

- (1) *Le système dynamique topologique  $(X, \phi)$  est uniquement ergodique, c'est-à-dire il admet une unique mesure de probabilité invariante ergodique.*
- (2) *Le système dynamique topologique  $(X, \phi)$  admet une unique mesure de probabilité invariante.*
- (3) *Pour tout  $f \in C^0(X; \mathbb{C})$ , il existe une constante  $L(f) \in \mathbb{C}$  telle que pour tout<sup>106</sup>  $x$  dans  $X$ , les moyennes de Birkhoff*

$$S_n f(x) = \frac{1}{n} \sum_{k=0}^{n-1} f \circ \phi^k(x)$$

*convergent quand  $n \rightarrow +\infty$  vers  $L(f)$ .*

- (4) *Il existe une partie  $E$  de  $C^0(X; \mathbb{C})$  engendrant un sous-espace vectoriel dense pour la topologie de la convergence uniforme sur  $C^0(X; \mathbb{C})$  tel que pour tout  $f \in E$ , et pour tout<sup>107</sup>  $x$  dans  $X$ , les moyennes de Birkhoff  $S_n f(x)$  convergent quand  $n \rightarrow +\infty$  vers une limite  $L(f)$  qui ne dépend pas de  $x$ .*

*Nous avons alors  $L(f) = \mu(f)$ , où  $\mu$  est l'unique mesure de probabilité invariante.*

**Démonstration.** Montrons que l'assertion (3) implique l'assertion (2). Soient  $\mu$  et  $\mu'$  deux mesures de probabilité  $\phi$ -invariantes (qui existent par le théorème de Krylov-Bogolyubov, voir la proposition 0.6). Le théorème ergodique de Birkhoff, appliqué à la mesure  $\mu$ , assure que pour tout  $f$  dans  $C^0(X; \mathbb{C})$ , la limite  $\tilde{f}(x)$  des moyennes de Birkhoff existe pour  $\mu$ -presque tout  $x \in X$  et vérifie  $\mu(\tilde{f}) = \mu(f)$ . Par hypothèse, l'application  $\tilde{f}$  est constante, donc égale à  $\mu(f)$ . En appliquant le même raisonnement à  $\mu'$ , nous avons donc  $\mu(f) = \mu'(f)$  et  $\mu = \mu'$ .

Réciproquement, montrons par contraposition que l'assertion (2) implique l'assertion (3). Soient  $x \in X$  et  $f \in C^0(X; \mathbb{C})$  tels que la suite  $S_n f(x)$  ne converge pas vers  $\mu(f)$ . Cette suite étant bornée, il existe une sous-suite  $(\tau(k))_{k \in \mathbb{N}}$  telle que la limite  $\lim_{k \rightarrow \infty} S_{\tau(k)} f(x)$  existe et est différente de  $\mu(f)$ . La compacité de  $\mathcal{M}(X)$  permet de supposer, quitte à extraire de nouveau une sous-suite, que la suite de mesures de probabilité

$$\mu_k = \frac{1}{\tau(k)} \sum_{i=0}^{\tau(k)-1} \Delta_{\phi^i(x)},$$

106. vraiment tout !

107. Voir la note de bas de page précédente !

où  $\Delta_y$  est la masse de Dirac unité en  $y$ , converge faiblement vers une mesure de probabilité  $\mu'$ . Par construction,  $\mu'$  est  $\phi$ -invariante et

$$\mu'(f) = \lim_{k \rightarrow \infty} S_{\tau(k)}f(x) \neq \mu(f) .$$

Donc  $\mu' \neq \mu$ .

L'équivalence entre les assertions (1) et (2) découle de la dernière assertion de la proposition 0.6 et de la proposition 2.1 (5).

L'équivalence entre les assertions (3) et (4) découle, en notant  $\| \cdot \|$  la norme uniforme sur  $C^0(X; \mathbb{C})$ , de la majoration uniforme, pour tous les  $f, g \in C^0(X; \mathbb{C})$ ,  $n \in \mathbb{N} \setminus \{0\}$  et  $x \in X$ ,

$$|S_n f(x) - S_n g(x)| \leq \|f - g\| .$$

□

Nous construirons dans les parties 11.3 et 11.4 de nombreuses mesures invariantes ergodiques pour les systèmes dynamiques symboliques et les endomorphismes linéaires du tore. Nous renvoyons à la proposition 4.3 (1) ci-dessous pour des exemples de systèmes de Kronecker uniquement ergodiques, ainsi qu'à l'exercice E.25.

### 4.3 Unique ergodicité des translations sur le tore

Le résultat ci-dessous généralise les propriétés déjà vues des rotations du cercle.

**Proposition 4.3.** Soient  $a = (a_1, \dots, a_N) \in \mathbb{R}^N$  et  $\alpha \in \mathbb{T}^N$  la classe modulo  $\mathbb{Z}^N$  de  $a$ . Notons  $\tau_\alpha : \mathbb{T}^N \rightarrow \mathbb{T}^N$  la translation  $\theta \mapsto \theta + \alpha$ . Les assertions suivantes sont équivalentes :

- (1)  $\tau_\alpha$  est uniquement ergodique ;
- (2)  $\tau_\alpha$  est ergodique pour  $\lambda = d\theta$  ;
- (3)  $\tau_\alpha$  admet une orbite dense ;
- (4) toutes les orbites de  $\tau_\alpha$  sont denses ;
- (5)  $\tau_\alpha$  est minimal, c'est-à-dire que le seul fermé invariant non vide de  $\mathbb{T}^N$  est  $\mathbb{T}^N$  ;
- (6) les réels  $1, a_1, \dots, a_N$  sont linéairement indépendants sur  $\mathbb{Q}$ .

**Démonstration.** (1)  $\Rightarrow$  (2) La mesure  $d\theta$  est invariante, donc ergodique par unicité.

(2)  $\Rightarrow$  (3) Cela résulte de l'exercice E.12.

(3)  $\Rightarrow$  (4) Si l'orbite d'un point  $\theta$  est dense dans  $\mathbb{T}^N$ , il en est de même de l'orbite de tous les autres points  $\theta'$ , car comme les translations  $\tau_\alpha$  et  $\tau_{\theta' - \theta}$  commutent, l'application continue  $\tau_{\theta' - \theta}$  envoie l'orbite de  $\theta$  par  $\tau_\alpha$  sur celle de  $\theta'$ .

(4)  $\Rightarrow$  (5) Ceci découle de l'équivalence des définitions d'un système dynamique minimal, voir la partie 7.4.

(5)  $\Rightarrow$  (6) Sinon, il existerait un élément non nul  $m$  de  $\mathbb{Z}^N$  tel que  $\langle m, \alpha \rangle = 0$ . L'ensemble des  $\theta$  dans  $\mathbb{T}^N$  tels que  $\langle m, \theta \rangle = 0$  dans  $\mathbb{R}/\mathbb{Z}$  serait donc un fermé non vide invariant par  $\tau_\alpha$ , différent de  $\mathbb{T}^N$ , contradiction.

(6)  $\Rightarrow$  (1) Soit  $\mu$  une mesure de probabilité borélienne invariante par  $\tau_\alpha$ . Nous avons donc les égalités

$$c_m(\mu) = \mu(e_{-m}) = \mu(e_{-m} \circ \tau_\alpha) = \mu(e^{-2i\pi\langle m, \alpha \rangle} e_{-m}) = e^{-2i\pi\langle m, \alpha \rangle} c_m(\mu).$$

Par hypothèse, lorsque  $m \neq 0$ , l'élément  $\langle m, \alpha \rangle$  est non nul dans  $\mathbb{R}/\mathbb{Z}$ , donc  $c_m(\mu) = 0$ . Nous en déduisons que  $c_m(\mu) = c_m(d\theta)$  pour tout  $m \in \mathbb{Z}^N$ . L'injectivité de la transformation de Fourier (voir la proposition 4.1 (2)) assure que  $\mu = d\theta$ .  $\square$

#### 4.4 Equirépartition de suites modulo 1 et unique ergodicité

Donnons quelques applications arithmétiques de l'unique ergodicité.

Pour tout réel  $x \in \mathbb{R}$ , notons  $[x] \in \mathbb{Z}$  sa *partie entière* inférieure et  $\{x\} = x - [x] \in [0, 1[$  sa *partie fractionnaire*. Une suite de réels  $(x_n)_{n \in \mathbb{N}}$  est dite *équirépartie modulo 1* si, pour tout intervalle  $[a, b]$  de  $[0, 1]$ , nous avons

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{Card} \{k \leq n / \{x_k\} \in [a, b]\} = b - a.$$

Le résultat suivant est très utile.

**Proposition 4.4. (Critère de Weyl <sup>104</sup>)** Soit  $(x_n)_{n \in \mathbb{N}}$  une suite de réels. Les assertions suivantes sont équivalentes.

- (1) La suite  $(x_n)_{n \in \mathbb{N}}$  est équirépartie modulo 1.
- (2) Pour tout  $f \in C^0([0, 1]; \mathbb{C})$ , nous avons

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\{x_k\}) = \int_0^1 f(t) dt.$$

- (3) Pour tout  $m$  dans  $\mathbb{Z} \setminus \{0\}$ , nous avons

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} e^{2i\pi m x_k} = 0.$$

**Démonstration.** Nous commençons la démonstration par le lemme suivant, dont nous conserverons les notations de son énoncé et de sa démonstration.

**Lemme 4.5.** Soit  $\mathcal{F}_b([0, 1]; \mathbb{C})$  l'espace vectoriel des fonctions bornées sur  $[0, 1]$ , muni de la norme  $\|\varphi\|_\infty = \sup_{t \in [0, 1]} |\varphi(t)|$ , et  $E$  l'ensemble des fonctions  $f \in \mathcal{F}_b([0, 1]; \mathbb{C})$  telles que  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\{x_k\}) = \int_0^1 f(t) dt$ . Alors  $E$  est un sous-espace vectoriel fermé dans  $\mathcal{F}_b([0, 1]; \mathbb{C})$ .

**Démonstration.** Soit  $f$  dans l'adhérence  $\overline{E}$  de  $E$  dans  $\mathcal{F}_b([0, 1]; \mathbb{C})$ . Pour tout  $\epsilon > 0$ , il existe donc  $g$  dans  $E$  tel que  $\|f - g\|_\infty \leq \epsilon$ . Notons  $\Sigma_n f = \frac{1}{n} \sum_{k=0}^{n-1} f(\{x_k\})$ . Pour tout  $n \geq 1$ , nous avons donc  $|\Sigma_n f - \Sigma_n g| \leq \epsilon$  et  $|\int_0^1 f - \int_0^1 g| \leq \epsilon$ . Comme  $\lim_{n \rightarrow \infty} |\Sigma_n g - \int_0^1 g| = 0$ , nous avons

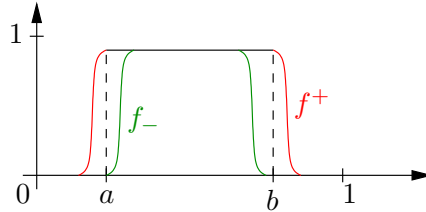
$$\limsup_{n \rightarrow \infty} \left| \Sigma_n f - \int_0^1 f \right| \leq 2\epsilon.$$



Ceci est vrai pour tout  $\epsilon$ . Donc  $\lim_{n \rightarrow \infty} |\Sigma_n f - \int_0^1 f| = 0$ . D'où  $f \in E$ , et  $E$  est fermé.  $\square$

Montrons tout d'abord que (1) implique (2) implique (3). Si la suite  $(x_n)_{n \in \mathbb{N}}$  est équirépartie modulo 1, alors par définition, les fonctions caractéristiques d'intervalles de  $[0, 1]$  sont dans l'ensemble  $E$  défini ci-dessus. Comme toute fonction continue sur  $[0, 1]$  est limite uniforme de fonctions en escalier, l'ensemble  $E$  contient  $C^0([0, 1]; \mathbb{C})$  par le lemme ci-dessus. En particulier, les fonctions  $e_m : x \mapsto e^{2i\pi m x}$  sont dans  $E$ . Nous concluons en remarquant que  $\int_0^1 e_m = 0$  pour  $m \neq 0$ .

Réciproquement, montrons que (3) implique (1). Pour tout  $m$  dans  $\mathbb{Z}$  (par l'hypothèse si  $m \geq 1$ ), la fonction  $e_m$  est dans  $E$ . Comme toute fonction 1-périodique, de classe  $C^2$ , est limite uniforme de sa série de Fourier, l'ensemble  $E$  contient les restrictions à  $[0, 1]$  de ces fonctions par le lemme ci-dessus. Soit  $[a, b]$  un intervalle contenu dans  $[0, 1]$ . Pour tout  $\epsilon > 0$ , nous pouvons trouver deux fonctions  $f_-, f_+$  dans  $E$  telles que  $f_- \leq \mathbb{1}_{[a,b]} \leq f_+$  et  $\int_0^1 f_+ - f_- \leq \epsilon$ .



Nous avons alors  $\Sigma_n f_- \leq \Sigma_n \mathbb{1}_{[a,b]} \leq \Sigma_n f_+$  et

$$b - a - \epsilon \leq \int_0^1 f_- \leq \liminf_{n \rightarrow \infty} \Sigma_n \mathbb{1}_{[a,b]} \leq \limsup_{n \rightarrow \infty} \Sigma_n \mathbb{1}_{[a,b]} \leq \int_0^1 f_+ \leq b - a + \epsilon.$$

Ceci est vrai pour tout  $\epsilon$ . Donc  $\lim_{n \rightarrow \infty} \Sigma_n \mathbb{1}_{[a,b]} = b - a$ .  $\square$

Le résultat suivant redémontre (par l'assertion (2) du critère de Weyl et l'assertion (3) de la proposition 4.2) l'unique ergodicité des rotations irrationnelles du cercle.

**Corollaire 4.6.** *Pour tout irrationnel  $\alpha$ , la suite  $(n\alpha)_{n \in \mathbb{N}}$  est équirépartie modulo 1.*

**Démonstration.** Ceci résulte du critère de Weyl, car pour tout  $m \neq 0$ , nous avons  $e^{2i\pi m \alpha} \neq 1$ , et donc

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} e^{2i\pi m k \alpha} \right| = \frac{1}{n} \left| \frac{1 - e^{2i\pi m n \alpha}}{1 - e^{2i\pi m \alpha}} \right| \leq \frac{2}{n |1 - e^{2i\pi m \alpha}|} \xrightarrow{n \rightarrow \infty} 0. \square$$

**Proposition 4.7.** *Soit  $\alpha$  un nombre irrationnel. L'homéomorphisme  $\phi$  du tore  $\mathbb{T}^N$  donné par*

$$\phi(\theta_1, \dots, \theta_N) = (\theta_1 + \alpha, \theta_2 + \theta_1, \dots, \theta_N + \theta_{N-1})$$

*est uniquement ergodique.*

**Démonstration.** Soit  $\mu$  une mesure de probabilité  $\phi$ -invariante sur  $\mathbb{T}^N$ . Nous voulons montrer que  $\mu = d\theta$ . Pour tout  $m$  dans  $\mathbb{Z}$ , le coefficient de Fourier  $c_m(\mu)$  vérifie, par invariance de  $\mu$ , l'égalité  $c_m(\mu) = \mu(e_{-m}) = \mu(e_{-m} \circ \phi)$ . Donc nous avons

$$c_{(m_1, \dots, m_N)}(\mu) = e^{-2i\pi m_1 \alpha} c_{(m_1 + m_2, m_2 + m_3, \dots, m_{N-1} + m_N, m_N)}(\mu). \quad (38)$$

Ces égalités à elles seules ne suffisent pas à montrer que  $c_m(\mu) = 1$  si  $m = 0$  et  $c_m(\mu) = 0$  sinon. Nous devons exploiter le fait que les nombres  $c_m(\mu)$  sont les coefficients de Fourier d'une mesure de probabilité. Notons

$$\langle \varphi, \psi \rangle_{\mathbb{L}^2(\mu)} = \int_{\mathbb{T}^N} \varphi(\theta) \overline{\psi}(\theta) d\mu(\theta)$$

le produit scalaire de  $\mathbb{L}^2(\mathbb{T}^N, \mu)$ .

Montrons par récurrence sur  $d$  que si  $m = (m_1, \dots, m_d, 0, \dots, 0) \neq 0$  alors  $c_m(\mu) = 0$ . Si  $m = (m_1, 0, \dots, 0) \neq 0$ , alors les égalités (38) assurent que  $c_m(\mu) = 0$ . Supposons le résultat vrai pour  $d - 1$ , et soit  $m = (m_1, \dots, m_d, 0, \dots, 0)$ .

Fixons  $M$  dans  $\mathbb{N} - \{0\}$ . Pour  $k = 1, \dots, M$ , les fonctions  $f_k = e_m \circ \phi^k$  sont de la forme  $\lambda_k e_{m(k)}$  avec  $\lambda_k$  un nombre complexe de module 1,  $m(k) = (m'_1, m'_2, \dots, m'_d, 0, \dots, 0)$  et  $m'_d = m_d$ . Par hypothèse de récurrence, nous avons donc

$$\langle f_k, f_{k'} \rangle_{\mathbb{L}^2(\mu)} = \lambda_k \overline{\lambda_{k'}} \mu(e_{m(k') - m(k)}) = \lambda_k \overline{\lambda_{k'}} c_{m(k) - m(k')}(\mu) = \begin{cases} 1 & \text{si } k = k' \\ 0 & \text{sinon.} \end{cases}$$

Autrement dit, la famille  $(f_k)_{k=1, \dots, M}$  est orthonormée. D'après le théorème de Pythagore, la fonction  $\psi = e_0$  vérifie

$$\sum_{k=1}^M |\langle f_k, \psi \rangle_{\mathbb{L}^2(\mu)}|^2 \leq \|\psi\|_{\mathbb{L}^2(\mu)}^2 = 1.$$

Nous calculons facilement que  $\langle f_k, \psi \rangle_{\mathbb{L}^2(\mu)} = \mu(e_{-m} \circ \phi^k) = \mu(e_{-m}) = c_m(\mu)$ . Donc, nous avons  $|c_m(\mu)|^2 \leq \frac{1}{M}$ . Comme ceci est vrai pour tout  $M$ , nous avons  $c_m(\mu) = 0$ .

Ceci prouve que  $\mu$  et  $d\theta$  ont les mêmes coefficients de Fourier. Donc  $\mu = d\theta$  par la proposition 4.1 (2).  $\square$

**Corollaire 4.8. (Théorème de Weyl)** *Soit  $P$  un polynôme non constant à coefficients réels dont le coefficient dominant est irrationnel. Alors la suite  $n \mapsto P(n)$  est équirépartie modulo 1.*

En particulier, pour tout irrationnel  $\alpha$ , la suite  $(n\alpha)_{n \in \mathbb{N}}$  est équirépartie modulo 1, comme vu dans le corollaire 4.6.

**Démonstration.** Notons  $N$  le degré de  $P$  et  $\frac{1}{N!}\alpha$  le coefficient dominant de  $P$ . Posons  $P_N = P$  et, par récurrence descendante,  $P_j(X) = P_{j+1}(X + 1) - P_{j+1}(X)$  pour  $j = 0, \dots, N - 1$ . Le polynôme  $P_j$  est de degré  $j$  et  $P_0 = \alpha$ . Posons  $\theta_n = (P_1(n), \dots, P_N(n)) \bmod \mathbb{Z}^N$ . Nous avons donc  $\phi(\theta_n) = \theta_{n+1}$  pour tout entier  $n$ . Comme  $\phi$  est uniquement ergodique, les points  $\theta_n = \phi^n(\theta_0)$  sont équirépartis sur  $\mathbb{T}^N$ , i.e. pour toute application continue  $f$  sur  $\mathbb{T}^N$ , nous avons

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\theta_k) = \int_{\mathbb{T}^N} f d\theta.$$

En choisissant  $f$  ne dépendant que de la dernière coordonnée, nous en déduisons que la suite  $P(n)$  est équirépartie modulo 1.  $\square$

## 4.5 Propriétés de mélange des endomorphismes linéaires du tore

Nous allons maintenant définir et étudier les transformations linéaires du tore, qui, tout en permettant une approche élémentaire, représentent des modèles frappants de dynamique riche et en particulier chaotique.

Soit  $M = (m_{ij})_{1 \leq i, j \leq N}$  une matrice  $N \times N$  à coefficients dans  $\mathbb{Z}$ , que nous identifions avec la transformation linéaire de  $\mathbb{R}^N$  de matrice  $M$  dans la base canonique. Puisque l'application linéaire  $M$  envoie  $\mathbb{Z}^N$  dans  $\mathbb{Z}^N$ , elle passe au quotient pour définir une application  $\phi_M : \mathbb{T}^N \rightarrow \mathbb{T}^N$ , qui envoie  $\theta = (\theta_i)_{1 \leq i \leq N}$  sur

$$\phi_M(\theta) = \left( \sum_{j=1}^N m_{ij} \theta_j \right)_{1 \leq i \leq N},$$

de sorte que le diagramme suivant commute, avec  $\mathcal{r} : x \mapsto \dot{x} = x \bmod \mathbb{Z}^N$  la projection canonique,

$$\begin{array}{ccc} \mathbb{R}^N & \xrightarrow{M} & \mathbb{R}^N \\ \mathcal{r} \downarrow & & \downarrow \mathcal{r} \\ \mathbb{T}^N & \xrightarrow{\phi_M} & \mathbb{T}^N \end{array} .$$

**Lemme 4.9.** *Supposons que le déterminant de  $M$  soit non nul. Alors*

- (1) *l'application  $\phi_M$  est surjective,*
- (2) *l'application  $\phi_M$  préserve la mesure de Haar  $\lambda = d\theta$  sur  $\mathbb{T}^N$ ,*
- (3) *pour tout  $\theta$  dans  $\mathbb{T}^N$ , nous avons  $\text{Card } \phi_M^{-1}(\theta) = |\det M|$ .*

**Démonstration.** (1) En effet,  $M$  et  $\mathcal{r}$  sont surjectives.

(2) Comme nous avons  $\phi_M \circ \tau_\theta = \tau_{\phi_M(\theta)} \circ \phi_M$  pour tout  $\theta$  dans  $\mathbb{T}^N$ , la mesure de probabilité image  $(\phi_M)_* \lambda$  est invariante par  $\tau_{\phi_M(\theta)}$ . Comme  $\phi_M$  est surjective, et puisque  $\lambda$  est la seule mesure de probabilité invariante par toutes les translations sur  $\mathbb{T}^N$ , nous avons  $(\phi_M)_* \lambda = \lambda$ .

(3) L'application  $\phi_M$  est un morphisme de groupes, donc toutes ses fibres ont le même cardinal. Comme l'application linéaire  $M$  est injective et puisque  $\mathcal{r}$  est un homéomorphisme local, le noyau  $\phi_M^{-1}(0)$  est un sous-groupe discret de  $\mathbb{T}^N$ , donc fini car  $\mathbb{T}^N$  est compact. Notons  $d = \text{Card } \phi_M^{-1}(0)$ ,  $A_\epsilon = ]-\epsilon, \epsilon[^N$  et  $B_\epsilon = \mathcal{r}(A_\epsilon)$  pour tout  $\epsilon$  dans  $]0, (2N \sup_{1 \leq i, j \leq N} |m_{ij}|)^{-1}[$ . Nous avons, par la commutativité du diagramme ci-dessus,

$$\mathcal{r}^{-1} \phi_M^{-1} \phi_M \mathcal{r}(A_\epsilon) = M^{-1} \mathcal{r}^{-1} \mathcal{r} M(A_\epsilon) = M^{-1}(MA_\epsilon + \mathbb{Z}^N) = A_\epsilon + M^{-1}\mathbb{Z}^N .$$

Puisque  $\phi_M^{-1}(0) = \mathcal{r}(M^{-1}\mathbb{Z}^N)$  et puisque  $\mathcal{r}$  est un morphisme de groupes, nous avons

$$\phi_M^{-1}(\phi_M(B_\epsilon)) = \bigcup_{\theta \in \phi_M^{-1}(0)} (\theta + B_\epsilon)$$

et cette réunion est une réunion disjointe par l'hypothèse sur  $\epsilon$ . Par invariance de  $\lambda$ , nous avons alors

$$\lambda(\phi_M(B_\epsilon)) = \lambda(\phi_M^{-1}(\phi_M(B_\epsilon))) = d \lambda(B_\epsilon) .$$

D'autre part, la formule de changement de variables dans  $\mathbb{R}^N$  donne

$$\lambda(\phi_M(B_\epsilon)) = |\det M| \lambda(B_\epsilon).$$

Donc  $d = |\det M|$ . □

En munissant  $\mathbb{T}^N$  de sa tribu borélienne, nous allons nous intéresser aux propriétés ergodiques du système dynamique mesuré  $(\mathbb{T}^N, \lambda, \phi_M)$ .

**Théorème 4.10.** *Soit  $\phi_M : \mathbb{T}^N \rightarrow \mathbb{T}^N$  la transformation du tore donnée par une matrice  $M$  à coefficients entiers de déterminant non nul. Les assertions suivantes sont équivalentes :*

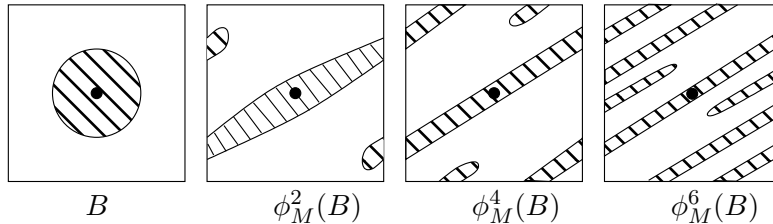
- (1)  $\phi_M$  est mélangeante pour la mesure de Haar  $\lambda = d\theta$ ,
- (2)  $\phi_M$  est ergodique pour la mesure de Haar  $\lambda = d\theta$ ,
- (3) aucune valeur propre (complexe) de  $M$  n'est racine de l'unité, ou, de manière équivalente,

$$\forall n \in \mathbb{N} - \{0\}, \quad \det(M^n - \text{id}) \neq 0.$$

De plus, si les valeurs propres (complexes) de  $M$  sont de module strictement supérieur à 1 (nous dirons que  $\phi_M$  est dilatante), si  $k = \lceil \frac{N+1}{2} \rceil$ , alors le système dynamique mesuré  $(\mathbb{T}^N, d\theta, \phi_M)$  est exponentiellement mélangeant sur l'espace vectoriel des applications de classe  $C^k$  sur  $\mathbb{T}^N$  muni de la norme de Sobolev  $\|\cdot\|_{W^{\infty,k}}$  : il existe des constantes  $c, \kappa > 0$  telles que pour tous les  $f, g : \mathbb{T}^N \rightarrow \mathbb{R}$  de classe  $C^k$ , et pour tout  $n \in \mathbb{N}$ , nous avons

$$\left| \int_X (f \circ \phi_M^n) g \, d\mu - \left( \int_X f \, d\mu \right) \left( \int_X g \, d\mu \right) \right| \leq c \|f\|_{\infty,k} \|g\|_{\infty,k} e^{-\kappa n}.$$

Voici quelques dessins avec la matrice  $M = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$  qui montrent bien le mélange.



**Démonstration.** Nous avons déjà vu que l'assertion (1) impliquait l'assertion (2) (voir la remarque (iii) ci-dessus). Nous allons donner une démonstration des autres implications du théorème 4.10 par analyse de Fourier. La remarque fondamentale est que pour tout  $m \in \mathbb{Z}^N$ , nous avons  $e_m \circ \phi_M = e_{m'}$  avec  $m' = {}^t M m$  l'image de  $m$  par la matrice transposée de  $M$ , par la définition des fonctions trigonométriques  $(e_m)_{m \in \mathbb{Z}^N}$  et le fait que la matrice adjointe, pour le produit scalaire usuel de  $\mathbb{R}^N$ , d'une matrice de  $\mathcal{M}_N(\mathbb{R})$  est sa matrice transposée.

Montrons que l'assertion (2) implique l'assertion (3). Par contraposée, supposons qu'il existe  $n \geq 1$  tel que  $\det(M^n - \text{id}) = 0$ , de sorte que 1 soit une valeur propre de la matrice à coefficients entiers (donc rationnels)  ${}^t M^n$ . Alors, soit  $m_0$  un élément non nul de  $\mathbb{Z}^N$  tel que  ${}^t M^n m_0 = m_0$  (commencer par le prendre dans  $\mathbb{Q}^N$  puis multiplier par le ppem des dénominateurs de ses composantes). Notons  $m_i = {}^t M^i m_0$  pour  $i = 0, \dots, n-1$ , qui est aussi un élément non nul de  $\mathbb{Z}^N$ . La fonction continue  $f = e_{m_0} + e_{m_1} + \dots + e_{m_{n-1}}$  est alors

invariante par  $\phi_M$  et non constante (car son produit scalaire  $\mathbb{L}^2$  avec  $e_0$  est nul). Donc  $\phi_M$  n'est pas ergodique.

Montrons que l'assertion (3) implique l'assertion (1). Par la proposition 3.1, il suffit de montrer que les coefficients de corrélation  $c_n(f, g)$  tendent vers 0 quand  $n \rightarrow +\infty$  pour  $f, g$  dans une partie de  $\mathbb{L}_0^2(\mathbb{T}^N, d\theta)$  engendrant un sous-espace vectoriel dense de  $\mathbb{L}_0^2(\mathbb{T}^N, d\theta)$ . Puisque  $(e_m)_{m \in \mathbb{Z}^N - \{0\}}$  est une base hilbertienne de  $\mathbb{L}_0^2(\mathbb{T}^N, d\theta)$ , nous pouvons donc supposer que  $f = e_m$  et  $g = e_\ell$  avec  $\ell, m$  dans  $\mathbb{Z}^N - \{0\}$ . En particulier  $\int_{\mathbb{T}^N} f d\theta = 0$ .

Pour tous les  $m' \in \mathbb{Z}^N$  et  $n \in \mathbb{N}$ , posons

$$m'_n = {}^t M^n m' \in \mathbb{Z}^N$$

(qui est non nul si et seulement si  $m'$  est non nul). Nous avons l'égalité

$$e_m \circ \phi_M^n = e_{m_n} . \quad (39)$$

Par hypothèse,  ${}^t M$  n'a pas de point périodique dans  $\mathbb{Z}^N - \{0\}$ . Donc la suite  $m_n$  prend des valeurs deux à deux distinctes. En particulier, il existe  $n_0$  dans  $\mathbb{N}$  tel que pour  $n \geq n_0$ , nous avons  $m_n \neq -\ell$ . Nous avons alors  $c_n(f, g) = \int_{\mathbb{T}^N} (f \circ \phi_M^n) g d\theta = 0$  par les relations d'orthogonalité de la proposition 4.1 (5), ce qui démontre la propriété de mélange.

Montrons la dernière assertion sur le mélange exponentiel. Supposons que les valeurs propres de  $M$  soient de module strictement supérieur à 1, et posons  $k = \lceil \frac{N+1}{2} \rceil$ . Montrons qu'il existe  $c, \kappa > 0$  tels que pour tous les  $f, g \in C^k(\mathbb{T}^N; \mathbb{C})$  et  $n \in \mathbb{N}$ , nous avons

$$|c_n(f, g)| \leq c \|f\|_{\infty, k} \|g\|_{\infty, k} e^{-\kappa n} .$$

Remarquons que pour tout  $k \in \mathbb{N}$ , nous avons

$$\|f - c_0(f)\|_{\infty, k} \leq \|f\|_{\infty, k} + |c_0(f)| \leq \|f\|_{\infty, k} + \|f\|_{\infty} \leq 2 \|f\|_{\infty, k} .$$

Puisque  $c_n(f, g) = 0$  si  $f$  est constant et par linéarité, quitte à remplacer  $f$  par  $f - c_0(f)$  et la constante  $c$  cherchée par  $4c$ , nous pouvons donc supposer que  $c_0(f) = 0$ . Par la transformation de Fourier inverse et sa convergence uniforme (voir la proposition 4.1 (3), en notant que  $\lceil \frac{N+1}{2} \rceil \geq 1$ , donc  $f$  et  $g$  sont de classe  $C^1$  au moins), par la formule (39) et les formules d'orthogonalité (37), nous avons, pour tout  $n \in \mathbb{N}$ ,

$$\begin{aligned} c_n(f, g) &= \int_{\mathbb{T}^N} (f \circ \phi_M^n) g d\theta \\ &= \int_{\mathbb{T}^N} \left( \sum_{m \in \mathbb{Z}^N - \{0\}} c_m(f) e_m \circ \phi_M^n(\theta) \right) \left( \sum_{m' \in \mathbb{Z}^N} c_{m'}(g) e_{m'}(\theta) \right) d\theta \\ &= \sum_{m, m' \in \mathbb{Z}^N - \{0\}} c_m(f) c_{m'}(g) \int_{\mathbb{T}^N} e_{m_n}(\theta) e_{m'}(\theta) d\theta \\ &= \sum_{m \in \mathbb{Z}^N - \{0\}} c_m(f) c_{-m_n}(g) . \end{aligned}$$

Pour tout  $m' \in \mathbb{Z}^N - \{0\}$ , posons  $\|m'\|_{\infty} = \max_{1 \leq j \leq N} |m'(j)|$  où  $m'(j) \in \mathbb{Z}$  est la  $j$ -ème composante de  $m'$ . Pour tous les  $\ell \in \mathbb{N}$  et  $h \in C^\ell(\mathbb{T}^N; \mathbb{C})$ , par la proposition 4.1 (4) et par

réurrence en considérant la dérivée partielle de  $h$  par rapport à une variable  $j \in \{1, \dots, N\}$  telle que  $\|m'\|_\infty = |m'(j)|$ , et puisque  $2\pi \geq 1$ , nous avons

$$|c_{m'}(h)| \leq \frac{1}{\|m'\|_\infty^\ell} \|h\|_{\infty, \ell}. \quad (40)$$

Puisque les valeurs propres  $\lambda_1, \dots, \lambda_N$  de  $M$  (donc de  ${}^tM$ ) sont de module strictement supérieur à 1, en posant  $\lambda = \inf_{1 \leq i \leq N} |\lambda_i| > 1$ , il existe une constante  $c' > 0$  telle que pour tous les  $m \in \mathbb{Z}^N - \{0\}$  et  $n \in \mathbb{N}$ , nous avons

$$\|m_n\|_\infty = \|{}^tM^n m\|_\infty \geq \|m\|_\infty \inf_{x \in \mathbb{R}^N : \|x\|_\infty = 1} \|{}^tM^n x\|_\infty \geq c' \lambda^n \|m\|_\infty. \quad (41)$$

Puisque  $k = \lceil \frac{N+1}{2} \rceil$ , la série  $\sum_{p \in \mathbb{N} - \{0\}} \frac{1}{p^{2k-N+1}}$  converge. Remarquons que pour tout  $p \in \mathbb{N}$ , il y a au plus  $2N(2p+1)^{N-1}$  éléments  $m \in \mathbb{Z}^N$  tels que  $\|m\|_\infty = p$ . Nous avons donc, par les formules (40) et (41),

$$\begin{aligned} |c_n(f, g)| &\leq \sum_{m \in \mathbb{Z}^N - \{0\}} |c_m(f)| |c_{-m_n}(g)| \leq \sum_{m \in \mathbb{Z}^N - \{0\}} \frac{1}{\|m\|_\infty^k \|m_n\|_\infty^k} \|f\|_{\infty, k} \|g\|_{\infty, k} \\ &\leq \sum_{p \in \mathbb{N} - \{0\}} \sum_{m \in \mathbb{Z}^N - \{0\}, \|m\|_\infty = p} \frac{1}{\|m\|_\infty^k (c' \lambda^n \|m\|_\infty)^k} \|f\|_{\infty, k} \|g\|_{\infty, k} \\ &= \sum_{p \in \mathbb{N} - \{0\}} \frac{2N(2p+1)^{N-1}}{(c')^k p^{2k} \lambda^{kn}} \|f\|_{\infty, k} \|g\|_{\infty, k} = c e^{-\kappa n} \|f\|_{\infty, k} \|g\|_{\infty, k}, \end{aligned}$$

en posant  $c = \sum_{p \in \mathbb{N} - \{0\}} \frac{2N(2p+1)^{N-1}}{(c')^k p^{2k}} < \infty$  et  $\kappa = k \ln \lambda > 0$ .  $\square$

## 4.6 Exercices

### Exercice E.25.

- (1) Soient  $X$  un espace topologique métrisable compact non vide et  $\phi : X \rightarrow X$  une application continue uniquement ergodique, de mesure de probabilité invariante  $\mu$ . Soit  $G$  un groupe topologique métrisable compact,  $\mu_G$  sa mesure de Haar normalisée et  $g : X \rightarrow G$  une application continue. Considérons la transformation  $\Phi : X \times G \rightarrow X \times G$  définie par

$$\Phi : (x, y) \mapsto (\phi(x), g(x)y).$$

Montrer que  $(X \times G, \Phi)$  est une extension du système dynamique topologique  $(X, \phi)$  et que  $(X \times G, \mu \otimes \mu_G, \Phi)$  est une extension du système dynamique probabilisé  $(X, \mu, \phi)$ .

- (2) Montrer que si  $\Phi$  est ergodique pour  $\mu \otimes \mu_G$ , alors le système dynamique  $(X \times G, \Phi)$  est uniquement ergodique.
- (3) Soient  $n \in \mathbb{N} \setminus \{0, 1\}$ ,  $\alpha \in \mathbb{R}$  irrationnel, et  $A = (a_{ij})_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{Z})$  une matrice entière unipotente triangulaire supérieure telle que  $a_{k, k+1} \neq 0$  pour tout  $k \in \llbracket 1, n-1 \rrbracket$ . Montrer que l'application  $\phi : \mathbb{T}^n \rightarrow \mathbb{T}^n$  définie par

$$X = (x_1, \dots, x_n) \mapsto XA + (\alpha, 0, \dots, 0) \pmod{\mathbb{Z}^n}$$

est uniquement ergodique.

**Exercice E.26.** Soient  $N \in \mathbb{N} - \{0\}$  et  $M \in \mathcal{M}_N(\mathbb{Z})$  une matrice entière  $N \times N$  telle que  $M$  n'admette pas de valeur propre complexe de module inférieur ou égal à 1. Notons  $\gamma: y \mapsto \dot{y}$  la projection canonique de  $\mathbb{R}^N$  sur le tore  $\mathbb{T}^N$ . Notons  $\phi_M: \mathbb{T}^N \rightarrow \mathbb{T}^N$  l'application lisse  $\theta = (\theta_i)_{1 \leq i \leq N} \mapsto (\sum_{1 \leq j \leq N} m_{ij} \theta_j)_{1 \leq i \leq N}$ , définie par passage au quotient modulo  $\mathbb{Z}^N$  de l'application linéaire  $M: \mathbb{R}^N \rightarrow \mathbb{R}^N$ .

(1) Montrer que le système dynamique topologique à temps discret  $(\mathbb{T}^N, \phi_M)$  est non inversible, que l'ensemble  $\text{Per}(\phi_M)$  des points périodiques de  $\phi_M$  est dense dans le tore  $\mathbb{T}^N$ , que le système dynamique topologique  $(\mathbb{T}^N, \phi_M)$  est non errant, qu'il est positivement transitif, et qu'il est topologiquement mélangeant.

(2) Notons  $C^0(\mathbb{T}^N)$  l'espace de Banach des applications continues réelles sur  $\mathbb{T}^N$ , muni de la norme uniforme  $\|\cdot\|_\infty$ . Notons  $L_M: C^0(\mathbb{T}^N) \rightarrow C^0(\mathbb{T}^N)$  l'application linéaire continue (de norme au plus 1), appelée *opérateur de transfert*, définie, pour tous les  $f \in C^0(\mathbb{T}^N)$  et  $x \in \mathbb{T}^N$ , par

$$L_M f: x \mapsto \frac{1}{\text{Card}(\phi_M^{-1}(x))} \sum_{z \in \phi_M^{-1}(x)} f(z).$$

i) Pour tous les  $f \in C^0(\mathbb{T}^N)$  et  $\dot{y} \in \mathbb{T}^N$ , montrer que

$$L_M f(\dot{y}) = \frac{1}{|\det M|} \sum_{[i] \in \mathbb{Z}^N / M\mathbb{Z}^N} f \circ \gamma(M^{-1}(y + i)).$$

ii) Montrer qu'en notant  $\lambda = d\theta$  la mesure de Haar de  $\mathbb{T}^N$ , nous avons, pour tous les  $f, g \in C^0(\mathbb{T}^N)$ ,

$$\int_{\mathbb{T}^N} (f \circ \phi_M) g \, d\lambda = \int_{\mathbb{T}^N} f (L_M g) \, d\lambda. \quad (42)$$

iii) Nous munissons  $\mathcal{L}(\mathbb{R}^N)$  de la norme d'opérateur usuelle. Montrer qu'il existe  $c > 0$  et  $\lambda > 1$  tel que pour tous les  $f \in C^1(\mathbb{T}^N)$  et  $n \in \mathbb{N}$ , nous avons

$$\left| L_M^n f(x) - \int_{\mathbb{T}^N} f \, d\lambda \right| \leq \frac{c}{\lambda^n} \sup_{x \in \mathbb{T}^N} \|d_x f\|. \quad (43)$$

iv) En déduire, lorsque  $M$  n'admet pas de valeur propre complexe de module inférieur ou égal à 1, une nouvelle démonstration de la propriété du mélange de  $\phi_M$  pour la mesure de Haar du tore.

v) En déduire, lorsque  $M$  n'admet pas de valeur propre complexe de module inférieur ou égal à 1, que la transformation  $\phi_M$  est exponentiellement mélangeante pour la mesure de Haar du tore, sur l'espace vectoriel  $C^1(\mathbb{T}^N)$  muni la norme de Sobolev  $W^{1,\infty}$ , définie par  $\|f\|_{1,\infty} = \max\{\|f\|_\infty, \|df\|_\infty\}$ , où  $\|df\|_\infty = \sup_{x \in \mathbb{T}^N} \|d_x f\|$ .

## 4.7 Indications pour la résolution des exercices

**Correction de l'exercice E.25.** Montrons tout d'abord que la mesure  $\mu \otimes \mu_G$  est invariante par  $\phi$ . Il suffit par linéarité et densité de montrer cette invariance pour les fonctions produits  $u \times v: (x, y) \mapsto u(x)v(y)$  pour  $u \in C^0(X; \mathbb{C})$  et  $v \in C^0(G; \mathbb{C})$ . Or par

le théorème de Fubini, l'invariance de  $\mu_G$  par les translations à gauche, et l'invariance de  $\mu$  par  $\phi$ , nous avons

$$\begin{aligned}
& \int_{X \times G} (u \times v) \circ \Phi \, d(\mu \otimes \mu_G) = \int_X \int_G u(\phi(x))v(g(x)y) \, d\mu(x)d\mu_G(y) \\
&= \int_X u(\phi(x)) \left( \int_G v(g(x)y) \, d\mu_G(y) \right) d\mu(x) = \int_X u(\phi(x)) \left( \int_G v(y) \, d\mu_G(y) \right) d\mu(x) \\
&= \int_G v(y) \, d\mu_G(y) \int_X u(\phi(x)) \, d\mu(x) = \int_G v(y) \, d\mu_G(y) \int_X u(x) \, d\mu(x) \\
&= \int_{X \times G} (u \times v) \, d(\mu \otimes \mu_G) .
\end{aligned}$$

La première projection  $h : X \times G \rightarrow G$  est alors une semi-conjugaison entre ces systèmes, car elle continue, surjective, vérifie  $h_*(\mu \otimes \mu_G) = \mu$  car  $\mu_G$  est une mesure de probabilité, et vérifie par définition  $h \circ \Phi = \phi \circ h$ .

**Correction de l'exercice E.26.** (1) Par le lemme 4.9 (3), nous savons que les fibres  $\phi_M^{-1}(\theta)$  pour  $\theta$  dans  $\mathbb{T}^N$  sont de cardinal égal à  $|\det M| > 1$ , donc  $\phi_M$  n'est pas injective.

Soit  $y \in \mathbb{R}^N$ . Pour tout  $n \in \mathbb{N} - \{0\}$ , nous avons  $\phi_M^n(\dot{y}) = \dot{y}$  si et seulement si  $M^n y \in y + \mathbb{Z}^N$ . Comme les valeurs propres complexes de  $M$  ne sont pas des racines de l'unité, la matrice  $M^n - \text{id}$  est inversible. Puisque toutes les valeurs propres complexes de  $M$  sont de module  $> 1$ , les valeurs propres de  $(M^n - \text{id})^{-1}$  tendent vers 0 quand  $n$  tend vers l'infini. Le réseau uniforme  $(M^n - \text{id})^{-1}(\mathbb{Z}^N)$  du groupe localement compact  $\mathbb{R}^N$  (voir la partie 9 pour la terminologie, qui n'est pas nécessaire ici) devient donc de plus en plus dense quand  $n$  tend vers  $+\infty$ , car son parallélépipède fondamental  $(M^n - \text{id})^{-1}([0, 1]^N)$  est de diamètre qui tend vers 0. Donc  $\bigcup_{n \in \mathbb{N} - \{0\}} (M^n - \text{id})^{-1}(\mathbb{Z}^N)$  est une partie dense de  $\mathbb{R}^N$ , et son image dans  $\mathbb{T}^N$  par l'application uniformément continue surjective  $\mathfrak{r}$ , qui est égale à  $\text{Per}(\phi_M)$ , est aussi dense.

Vu les hypothèses sur les valeurs propres de  $M$ , nous savons par le théorème 4.10 que la transformation  $\phi_M$  du tore est mélangeante pour la mesure de Haar de  $\mathbb{T}^N$ , donc est topologiquement mélangeante par l'exercice E.50, donc est positivement transitive et non errante comme vu dans la partie 7.4.

(2) Soient  $y, z \in \mathbb{R}^N$ . Pour tout  $n \in \mathbb{N} - \{0\}$ , nous avons  $\phi_M^n(\dot{z}) = \dot{y}$  si et seulement si  $M^n z \in y + \mathbb{Z}^N$ . Notons que  $M\mathbb{Z}^N$  est un sous-groupe d'indice fini (égal à  $|\det M|$ ) de  $\mathbb{Z}^N$ . Donc  $\phi_M^{-n}(\dot{y}) = \mathfrak{r}(\{M^{-n}(y + i) : [i] \in \mathbb{Z}^N / M^n \mathbb{Z}^N\})$ .

i) L'assertion i) découle alors du lemme 4.9 (3).

ii) Lorsque  $i$  parcourt  $\mathbb{Z}^N$ , les compacts  $M^{-1}([0, 1]^N + i)$  recouvrent  $\mathbb{R}^N$ . Ils sont d'intérieurs deux à deux disjoints, et de frontière de mesure de Lebesgue nulle. L'application  $\tau_i \circ M : M^{-1}([0, 1]^N + i) \rightarrow [0, 1]^N$  est un  $C^\infty$ -difféomorphisme (affine) de jacobien (constant) égal à  $|\det M|$ , et d'inverse  $y \mapsto M^{-1}(y + i)$ .

De plus, lorsque  $i$  parcourt un système de représentants des classes  $[i] \in \mathbb{Z}^N / M\mathbb{Z}^N$ , les images de ces compacts  $M^{-1}([0, 1]^N + i)$  dans  $\mathbb{T}^N$  sont d'intérieurs deux à deux disjoints, et recouvrent  $\mathbb{T}^N$ . Pour tout  $f \in C^0(\mathbb{T}^N)$ , notons  $\tilde{f} = f \circ \mathfrak{r}$ , qui est  $\mathbb{Z}^N$ -périodique (c'est-à-dire  $\tilde{f}(y + i) = \tilde{f}(y)$  pour tous les  $i \in \mathbb{Z}^N$  et  $y \in \mathbb{R}^N$ ) et continue sur  $\mathbb{R}^N$ . Alors pour tous les  $f \in C^0(\mathbb{T}^N)$ , en effectuant les changements de variables ci-dessus, par l'assertion



i), nous avons

$$\begin{aligned}
\int_{\mathbb{T}^N} (f \circ \phi_M) g \, d\lambda &= \sum_{[i] \in \mathbb{Z}^N / M\mathbb{Z}^N} \int_{M^{-1}([0,1]^N + i)} \tilde{f}(Mx) \tilde{g}(x) \, dx \\
&= \sum_{[i] \in \mathbb{Z}^N / M\mathbb{Z}^N} \int_{[0,1]^N + i} \tilde{f}(y + i) \tilde{g}(M^{-1}(y + i)) \frac{1}{|\det M|} \, dy \\
&= \int_{\mathbb{T}^N} f(L_M g) \, d\lambda .
\end{aligned}$$

## 5 Entropie mesurée

L'entropie ou *entropie mesurée*<sup>108</sup> d'un système dynamique mesuré  $(X, \mathcal{B}, \mu, \phi)$  est un invariant numérique (positif ou nul) de conjugaison mesurée d'un tel système, qui mesure le « désordre créé par sa transformation » d'un point de vue de sa mesure. Son origine remonte à l'entropie de Shannon<sup>104</sup> en théorie de l'information. Elle a été développée par Kolmogorov et Sinai<sup>104</sup> pour les systèmes dynamiques dans les années 1950, précisément dans le but de distinguer à conjugaison mesurée près des systèmes dynamiques, même si c'est loin d'être un invariant complet.

Nous définissons tout d'abord l'entropie  $H_\mu(\alpha)$  d'une partition  $\alpha$  de  $X$  comme l'information moyenne issue de la connaissance de la partie de  $\alpha$  dans laquelle se trouve un point de  $X$ . Toute partition de  $X$  permet de pister une trajectoire de  $\phi$  par la suite des parties visitées par la trajectoire. L'entropie  $h_\mu(\phi, \alpha)$  de  $\phi$  relativement à  $\alpha$  est alors le taux de croissance en  $n$  de l'information moyenne issue de la connaissance de la suite des parties de  $\alpha$  dans lesquelles se trouvent les  $n$  premiers points des orbites de  $\phi$ . L'entropie  $h_\mu(\phi)$  est alors la borne supérieure des entropies de  $f$  relativement aux partitions de  $X$ .

### 5.1 Information et entropie d'une partition

Commençons par une présentation heuristique de ces concepts. Lorsqu'on découpe un espace en  $p = 2^n$  parties de « tailles » égales (par un procédé de dichotomie par exemple), savoir dans quelle partie un point de  $X$  se trouve est une information qui peut s'exprimer à l'aide de  $n$  chiffres 0 ou 1. Nous dirons que l'information de cette partition est cet entier  $n = \log_2(p)$ , ou, plutôt, le réel  $\ln p$  qui lui est proportionnel. Lorsque le découpage n'est plus en parts égales, l'information est une fonction de  $x$  égale à la valeur absolue du logarithme de la « taille » de la partie dans laquelle se trouve  $x$ . La valeur moyenne de cette information est alors l'entropie de la partition.

Soit  $(X, \mathcal{B}, \mu)$  un espace de probabilité. Nous appellerons *partition mesurable* de  $(X, \mathcal{B})$  une famille finie  $\alpha = \{A_1, \dots, A_p\}$  de parties mesurables non vides de  $X$ , appelées les *événements* (ou parfois *atomes*) de  $\alpha$ , telles que, pour  $i \neq j$ , nous ayons  $A_i \cap A_j = \emptyset$  et  $A_1 \cup \dots \cup A_p = X$ . La *partition triviale* est la partition mesurable  $\alpha = \{X\}$  n'ayant qu'un seul atome. Pour tout  $x \in X$ , l'unique élément  $\alpha(x)$  de  $\alpha$  qui contient  $x$  est appelé l'*atome* de  $x$  pour  $\alpha$ , et son *information* est  $I_\alpha(x) = -\ln \mu(\alpha(x))$ . Dans la suite de la partie 5, les symboles  $\alpha, \beta, \gamma, \dots$  désignerons des partitions mesurables.

Nous pouvons aussi admettre des familles dénombrables, avec des intersections deux à deux de mesure nulle, et de réunion de mesure totale dans  $X$ . C'est parfois plus pratique pour des calculs explicites. Cela ne change pas les définitions et résultats qui suivent (voir par exemple [PY, Chap. 8] [KH, Part 1, 4.3]), sauf que

- les égalités d'ensemble sont des égalités modulo ensemble de mesure nulle, et en particulier, une partition est triviale si elle admet un (et alors un seul) élément de mesure total et les atomes sont alors bien définis  $\mu$ -presque partout, ce qui suffit dans la suite,
- l'entropie de certaines partitions dénombrables peut être infinie, et qu'il faut parfois exclure les partitions dénombrables d'entropie infinie ou demander une condition de non indétermination de type  $+\infty - \infty$  dans certains énoncés.

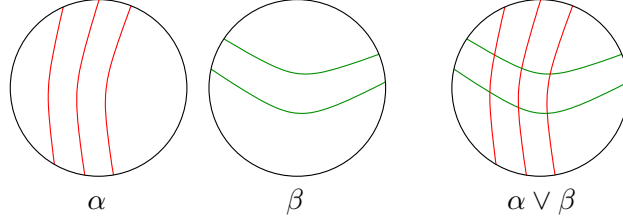
---

108. Une terminologie usuelle, mais peu appropriée car il n'y a pas de distance impliquée (ni nécessairement de distance sous-jacente), est celle d'*entropie métrique*.

L'avantage de la définition de partition mesurable que nous utilisons est qu'elle est indépendante de la mesure  $\mu$ , et permet donc de faire varier cette dernière. Voir aussi la note de bas de page 113 pour passer de dénombrable à fini.

**Partitions jointes.** Appelons *partition jointe* de  $\alpha$  et  $\beta$  la partition mesurable

$$\alpha \vee \beta = \{A \cap B : A \in \alpha, B \in \beta, A \cap B \neq \emptyset\}.$$



L'opérateur binaire  $\vee$  est une loi de composition interne sur l'ensemble  $\mathcal{P} = \mathcal{P}_{X, \mathcal{B}}$  des partitions mesurables de  $X$ , qui est associative<sup>109</sup>, commutative<sup>110</sup> et d'élément neutre la partition triviale<sup>111</sup>. Pour  $\alpha_1, \dots, \alpha_n \in \mathcal{P}$ , notons

$$\bigvee_{i=1}^n \alpha_i = \alpha_1 \vee \dots \vee \alpha_n = \{A_1 \cap \dots \cap A_n : \forall i = 1, \dots, n, A_i \in \alpha_i, A_1 \cap \dots \cap A_n \neq \emptyset\}.$$

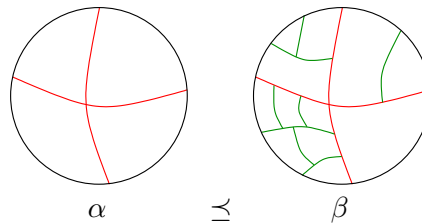
Remarquons que pour tout  $x \in X$ , nous avons

$$\left( \bigvee_{i=1}^n \alpha_i \right)(x) = \bigcap_{i=1}^n \alpha_i(x).$$

Si  $(\alpha_n)_{n \in \mathbb{N}}$  est une suite de partitions mesurables de  $(X, \mathcal{B})$ , nous noterons  $\sigma(\bigvee_{n \in \mathbb{N}} \alpha_n)$  la  $\sigma$ -algèbre engendrée par  $\bigcup_{n \in \mathbb{N}} (\bigvee_{i=0}^n \alpha_i)$ . Nous dirons que  $(\alpha_n)_{n \in \mathbb{N}}$  est *asymptotiquement génératrice* si  $\sigma(\bigvee_{n \in \mathbb{N}} \alpha_n) = \mathcal{B}$ .

Par exemple, si  $X = \mathbb{R}$  est muni de sa  $\sigma$ -algèbre borélienne, alors la suite de partitions mesurables  $(\alpha_n = \{[\frac{k}{n}, \frac{k+1}{n}] : k \in \mathbb{Z}\})_{n \in \mathbb{N}}$  est asymptotiquement génératrice. Si  $\mathcal{A}$  est un alphabet non vide, si  $X = \mathcal{A}^{\mathbb{N}}$  (respectivement  $X = \mathcal{A}^{\mathbb{Z}}$ ) est muni de la  $\sigma$ -algèbre produit, alors la suite de partitions mesurables  $(\alpha_n = \{[a_0, \dots, a_n] : a_0, \dots, a_n \in \mathcal{A}\})_{n \in \mathbb{N}}$  (respectivement  $(\alpha_n = \{[a_{-n}, \dots, a_n] : a_{-n}, \dots, a_n \in \mathcal{A}\})_{n \in \mathbb{N}}$ ) est asymptotiquement génératrice.

**Raffinement de partitions.** Nous dirons que  $\alpha$  est *moins fine* que  $\beta$  (ou que  $\beta$  est *plus fine* que  $\alpha$ ), et nous noterons  $\alpha \preceq \beta$ , si tout évènement de  $\alpha$  est réunion d'évènements de  $\beta$ , ou, de manière équivalente si tout atome de  $\beta$  est contenu dans un atome de  $\alpha$ . Nous dirons aussi que  $\beta$  est un *raffinement* de  $\alpha$ .



109. pour tous les  $\alpha, \beta, \gamma \in \mathcal{P}$ , nous avons  $(\alpha \vee \beta) \vee \gamma = \alpha \vee (\beta \vee \gamma)$

110. pour tous les  $\alpha, \beta \in \mathcal{P}$ , nous avons  $\alpha \vee \beta = \beta \vee \alpha$

111. pour tous les  $\alpha \in \mathcal{P}$ , nous avons  $\alpha \vee \{X\} = \alpha$

La relation  $\preceq$  est une relation d'ordre sur l'ensemble  $\mathcal{P}$  des partitions mesurables de  $(X, \mathcal{B})$ . Notons que

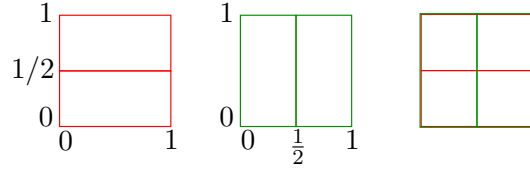
- la partition triviale est l'unique minimum de cette relation d'ordre : si  $\alpha = \{X\}$ , alors pour tout  $\beta \in \mathcal{P}$ , nous avons  $\alpha \preceq \beta$  et si  $\beta \preceq \alpha$ , alors  $\beta = \alpha$ .
- nous avons  $\alpha \preceq \beta$  si et seulement si  $\alpha \vee \beta = \beta$ ,
- $\alpha \vee \beta$  est la partition la moins fine qui est plus fine à la fois que  $\alpha$  et que  $\beta$ ;
- si  $\alpha \preceq \beta$  et  $\alpha' \preceq \beta'$  alors  $(\alpha \vee \alpha') \preceq (\beta \vee \beta')$ .

$$(44)$$

**Indépendance de partitions.** Deux partitions mesurables  $\alpha$  et  $\beta$  de  $(X, \mathcal{B})$  sont dites *indépendantes* pour  $\mu$ , et nous noterons  $\alpha \perp_\mu \beta$  (ou  $\alpha \perp \beta$  si  $\mu$  est sous-entendue), si pour tous les  $A \in \alpha$  et  $B \in \beta$ , nous avons

$$\mu(A \cap B) = \mu(A) \mu(B).$$

Par exemple, pour la mesure de Lebesgue sur le carré  $X = [0, 1[ \times [0, 1[$ , la partition en deux rectangles horizontaux  $\{[0, 1[ \times [0, \frac{1}{2}[, [0, 1[ \times [\frac{1}{2}, 1[ \}$  et celle en deux rectangles verticaux  $\{[0, \frac{1}{2}[ \times [0, 1[, [\frac{1}{2}, 1[ \times [0, 1[ \}$  sont indépendantes.



Notons que si  $\alpha$  et  $\beta$  sont des partitions mesurables de  $(X, \mathcal{B})$  indépendantes pour  $\mu$ , puisque  $(\alpha \vee \beta)(x) = \alpha(x) \cap \beta(x)$  pour tout  $x \in X$  et puisque  $\ln$  est un morphisme de groupes de  $([0, +\infty[ \times)$  dans  $(\mathbb{R}, +)$ , nous avons

$$\forall x \in X, \quad I_{\alpha \vee \beta}(x) = I_\alpha(x) + I_\beta(x).$$

Ainsi la fonction d'information  $I_\alpha : X \rightarrow [0, +\infty]$  définie par  $x \mapsto -\ln(\mu(\alpha(x)))$  vérifie les propriétés heuristiques suivantes.

- L'information est une quantité positive ou nulle (c'est la raison de la présence du signe  $-$  dans la définition).
- L'information apportée par l'appartenance à un évènement de la partition triviale est nulle : pour tout  $x \in X$ , nous avons  $I_{\{X\}}(x) = 0$ .
- Plus un évènement est improbable, plus l'information apportée par l'appartenance à cet évènement est grande : si  $\alpha \preceq \beta$ , alors  $I_\alpha \leq I_\beta$ .
- L'information apportée par l'appartenance à deux évènements indépendants est additive (c'est la raison de la présence du logarithme dans la définition, outre l'heuristique de processus de dichotomie donné au début de cette partie) : si  $\alpha \perp \beta$ , alors  $I_{\alpha \vee \beta} = I_\alpha + I_\beta$ .

**Égalité presque partout de partitions.** Notons  $A \Delta B = (A \setminus B) \sqcup (B \setminus A)$  la différence symétrique entre des ensembles  $A$  et  $B$ . Soient  $\alpha, \alpha', \beta, \beta'$  des partitions mesurables de  $(X, \mathcal{B})$ .

Nous dirons que  $\alpha$  et  $\beta$  sont *égales  $\mu$ -presque partout* (ou égales presque partout lorsque la mesure  $\mu$  est sous-entendue), et nous noterons  $\alpha = \beta \bmod_\mu 0$  (ou  $\alpha = \beta \bmod 0$  lorsque la mesure  $\mu$  est sous-entendue) si tout élément de  $\alpha$  de mesure non nulle est égal  $\mu$ -presque partout à une partie de  $\beta$ , ou de manière équivalente, s'il existe  $n \in \mathbb{N}$ , des parties  $\alpha' = \{A_1, \dots, A_n\}$  de  $\alpha$  et  $\beta' = \{B_1, \dots, B_n\}$  de  $\beta$  telles que

$$\forall i \in \llbracket 1, n \rrbracket, \quad \mu(A_i \Delta B_i) = 0$$

et  $\mu(A) = 0$  pour tout  $A \in \alpha \setminus \alpha'$  et  $\mu(B) = 0$  pour tout  $B \in \beta \setminus \beta'$ . La relation « être égales  $\mu$ -presque partout » est une relation d'équivalence sur l'ensemble  $\mathcal{P}$  des partitions mesurables de  $(X, \mathcal{B})$ . Nous noterons  $\mathcal{P}_{\text{mod}_\mu 0}$  l'ensemble des classes d'équivalence.

Remarquons que si  $\alpha = \beta \text{ mod } 0$  et  $\alpha' = \beta' \text{ mod } 0$ , alors nous avons  $\alpha \perp \alpha'$  si et seulement si  $\beta \perp \beta'$ . Notons que  $\alpha = \{X\} \text{ mod } 0$  si et seulement si  $\alpha$  contient un atome de masse totale.

De même, nous dirons que  $\alpha$  est  $\mu$ -presque partout moins fine que  $\beta$  (ou presque partout moins fine lorsque la mesure  $\mu$  est sous-entendue), et nous noterons  $\alpha \preceq \beta \text{ mod}_\mu 0$  (ou  $\alpha \preceq \beta \text{ mod } 0$  lorsque la mesure  $\mu$  est sous-entendue) si tout élément de  $\alpha$  de mesure non nulle est égal  $\mu$ -presque partout à une réunion de parties de  $\beta$ , ou de manière équivalente, si pour tout évènement  $A \in \alpha$ , il existe une partie  $\beta_A$  de  $\beta$  telle que

$$\mu(A \setminus (\cup \beta_A)) = 0 .$$

De manière équivalente,  $\alpha \preceq \beta \text{ mod } 0$  si et seulement si pour  $\mu$ -presque tout  $x \in X$ , l'atome  $\beta(x)$  de  $x$  pour  $\beta$  est presque partout contenu dans l'atome  $\alpha(x)$  de  $x$  pour  $\alpha$ . Remarquons que nous avons simultanément  $\alpha \preceq \beta \text{ mod } 0$  et  $\beta \preceq \alpha \text{ mod } 0$  si et seulement si  $\alpha = \beta \text{ mod } 0$ .

**Partitions images réciproques.** Pour toutes les partitions mesurables  $\alpha'$  et  $\beta'$  d'un espace mesurable  $(X', \mathcal{B}')$ , et pour toute application mesurable  $\phi : X \rightarrow X'$ , notons  $\phi^{-1}\alpha'$  la partition mesurable de  $X$ , appelée *partition image réciproque* de  $\alpha'$  par  $\phi$ , définie par

$$\phi^{-1}\alpha' = \{\phi^{-1}(A) : A \in \alpha'\} .$$

Nous avons

$$\phi^{-1}(\alpha' \vee \beta') = (\phi^{-1}\alpha') \vee (\phi^{-1}\beta') \tag{45}$$

et

$$\text{si } \alpha' \preceq \beta', \text{ alors } \phi^{-1}\alpha' \preceq \phi^{-1}\beta' . \tag{46}$$

Soit  $\mu' = \phi_*\mu$ . Nous avons de plus

$$\text{si } \alpha' \perp_{\mu'} \beta', \text{ alors } \phi^{-1}\alpha' \perp_\mu \phi^{-1}\beta'$$

et

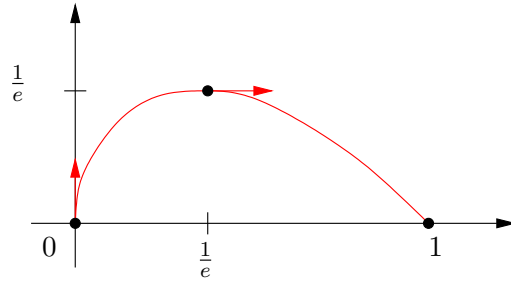
$$\text{si } \alpha' = \beta' \text{ mod}_{\mu'} 0, \text{ alors } \phi^{-1}\alpha' = \phi^{-1}\beta' \text{ mod}_\mu .$$

Notons que l'application

$$\varphi : t \mapsto -t \ln t$$

de l'intervalle  $[0, 1]$  dans lui-même (avec la convention justifiée pour obtenir la continuité que  $\varphi(0) = 0$ ) est une application continue, positive et strictement concave<sup>112</sup> sur  $[0, 1]$ , s'annulant exactement en 0 et en 1, de graphe suivant.

112. c'est-à-dire que pour tous les entiers  $k \geq 1$ , les points  $x_1, \dots, x_k$  de  $[0, 1]$  et les paramètres barycentriques  $s_1, \dots, s_k \in [0, 1]$  tels que  $\sum_{i=1}^k s_i = 1$ , nous avons  $\varphi(\sum_{i=1}^k s_i x_i) \geq \sum_{i=1}^k s_i \varphi(x_i)$  avec égalité si et seulement si  $x_1 = \dots = x_k$



L'entropie d'une partition mesurable  $\alpha$  de  $(X, \mathcal{B})$  pour la mesure  $\mu$  est

$$H_\mu(\alpha) = \sum_{A \in \alpha} \varphi(\mu(A)) ,$$

qui appartient à  $[0, +\infty[$  si  $\alpha$  est finie, et à  $[0, +\infty]$  si  $\alpha$  est dénombrable<sup>113</sup> Nous avons plus précisément  $H_\mu(\alpha) \in [-\ln(\sup_{A \in \alpha} \mu(A)), -\ln(\inf_{A \in \alpha} \mu(A))]$  et  $H_\mu(\alpha) = 0$  si et seulement si  $\alpha$  contient un atome de masse totale pour  $\mu$ .

L'entropie relative<sup>114</sup> d'une partition mesurable  $\alpha$  par rapport à une partition mesurable  $\beta$  est, avec la convention que  $\varphi\left(\frac{\mu(A \cap B)}{\mu(B)}\right)\mu(B) = 0$  si  $\mu(B) = 0$ ,

$$H_\mu(\alpha | \beta) = \sum_{A \in \alpha, B \in \beta} \varphi\left(\frac{\mu(A \cap B)}{\mu(B)}\right)\mu(B) .$$

Elle est à valeur dans  $[0, +\infty]$ , et est finie si  $\alpha$  et  $\beta$  sont finies. Notons qu'en prenant  $\beta = \{X\}$  dans cette formule, nous avons<sup>115</sup>

$$H_\mu(\alpha | \{X\}) = H_\mu(\alpha) .$$

L'information relative  $I_{\alpha|\beta}$  est nulle  $\mu$ -presque partout si et seulement si l'entropie relative  $H_\mu(\alpha | \beta)$  est nulle si et seulement si  $\alpha \preceq \beta \bmod_\mu 0$  si et seulement si  $\alpha \vee \beta = \beta$   $\mu$ -presque partout. Si  $\alpha = \alpha' \bmod_\mu 0$  et  $\beta = \beta' \bmod_\mu 0$ , alors

$$H_\mu(\alpha | \beta) = H_\mu(\alpha' | \beta') \quad \text{et} \quad H_\mu(\alpha) = H_\mu(\alpha') . \quad (47)$$

Nous noterons parfois  $H(\alpha) = H_\mu(\alpha)$  et  $H(\alpha | \beta) = H_\mu(\alpha | \beta)$  lorsque la mesure  $\mu$  est sous-entendue.

Les entropies et entropies relatives des partitions vérifient les propriétés suivantes.

**Proposition 5.1.** *Nous avons les propriétés suivantes.*

113. Un argument qui permet souvent de se ramener au cas des partitions finies est le suivant. Si  $\alpha = \{A_n : n \in \mathbb{N}\}$  est une partition mesurable dénombrable telle que  $H_\mu(\alpha) < +\infty$ , alors pour tout  $n \in \mathbb{N}$ , posons  $\alpha_n = \{A_0, \dots, A_n, \bigcup_{k>n} A_k\}$ , qui est une partition mesurable finie. Nous avons  $\alpha_n \preceq \alpha$  pour tout  $n \in \mathbb{N}$  et  $\lim_{n \rightarrow +\infty} H_\mu(\alpha_n) = H_\mu(\alpha)$ . En effet, nous avons  $\lim_{n \rightarrow +\infty} \mu(\bigcup_{k>n} A_k) = 0$  car la mesure est finie, et le reste de la série convergente définissant  $H_\mu(\alpha)$  tend vers 0.

114. On dit aussi l'entropie conditionnelle de  $\alpha$  par rapport à  $\beta$  ou l'entropie de  $\alpha$  conditionnellement à  $\beta$ , car  $\frac{\mu(A \cap B)}{\mu(B)}$  est l'entropie conditionnelle  $\mu(A | B)$  de  $A$  sachant  $B$  si  $\mu(B) \neq 0$ .

115. conformément à l'heuristique que l'information apportée par l'appartenance à un évènement de la partition triviale est nulle

(1) **(Majoration cardinale de l'entropie des partitions)** Pour toute partition mesurable  $\alpha$  de  $X$ , nous avons l'inégalité  $H_\mu(\alpha) \leq \ln \text{Card}(\alpha)$ , avec égalité si et seulement si, pour tout  $A$  dans  $\alpha$ , nous avons  $\mu(A) = \frac{1}{\text{Card}(\alpha)}$ .<sup>116</sup>

(2) **(Naturalité de l'entropie des partitions)** Pour tout espace mesurable  $X'$ , pour toute application mesurable  $\phi : X \rightarrow X'$  et toutes les partitions mesurables  $\alpha', \beta'$  de  $X'$ , nous avons

$$H_\mu(\phi^{-1}\alpha' \mid \phi^{-1}\beta') = H_{\phi_*\mu}(\alpha' \mid \beta') .$$

En particulier,  $H_\mu(\phi^{-1}\alpha') = H_{\phi_*\mu}(\alpha')$ .

(3) **(Décomposition de l'entropie des partitions)** Pour toutes les partitions mesurables  $\alpha, \beta$  et  $\gamma$  de  $X$ , nous avons

$$H_\mu(\alpha \vee \beta \mid \gamma) = H_\mu(\alpha \mid \beta \vee \gamma) + H_\mu(\beta \mid \gamma) .$$

En particulier,  $H_\mu(\alpha \vee \beta) = H_\mu(\alpha \mid \beta) + H_\mu(\beta)$ .

(4) **(Monotonie de l'entropie des partitions)** Soient  $\alpha, \beta$  et  $\gamma$  des partitions mesurables de  $X$ .

$$(i) \quad \text{Si } \alpha \preceq \beta \text{ alors } H_\mu(\alpha \mid \gamma) \leq H_\mu(\beta \mid \gamma) ,$$

et en particulier  $H_\mu(\alpha) \leq H_\mu(\beta)$  avec égalité si et seulement si  $\alpha = \beta \bmod_\mu 0$ .

$$(ii) \quad \text{Si } \gamma \preceq \beta \text{ alors } H_\mu(\alpha \mid \beta) \leq H_\mu(\alpha \mid \gamma) ,$$

et en particulier  $H_\mu(\alpha \mid \beta) \leq H_\mu(\alpha)$  avec égalité si et seulement si  $\alpha \perp \beta$ .

$$(iii) \quad \text{Nous avons}^{117} \quad H_\mu(\alpha) - H_\mu(\beta) \leq H_\mu(\alpha \mid \beta) .$$

(5) **(Sous-additivité de l'entropie des partitions)** Pour toutes les partitions mesurables  $\alpha, \beta, \gamma, \alpha_1, \dots, \alpha_n$  de  $X$ , nous avons

$$H_\mu(\alpha \vee \beta \mid \gamma) \leq H_\mu(\alpha \mid \gamma) + H_\mu(\beta \mid \gamma)$$

et en particulier  $H_\mu(\alpha \vee \beta) \leq H_\mu(\alpha) + H_\mu(\beta)$  avec égalité si et seulement si  $\alpha \perp \beta$ , et  $H_\mu(\bigvee_{i=1}^n \alpha_i) \leq \sum_{i=1}^n H_\mu(\alpha_i)$ .

(6) **(Concavité de l'entropie des partitions en la mesure)** Soient  $(X', \mathcal{B}')$  un espace mesurable et  $\alpha$  une partition mesurable de  $X'$ . Pour toutes les mesures de probabilités  $\nu_1, \dots, \nu_k$  sur  $X'$  et tous les  $s_1, \dots, s_k \in [0, 1]$  tels que  $\sum_{i=1}^k s_i = 1$  (appelés des paramètres barycentriques), nous avons

$$\sum_{i=1}^k s_i H_{\nu_i}(\alpha) \leq H_{\sum_{i=1}^k s_i \nu_i}(\alpha) \leq \sum_{i=1}^k s_i H_{\nu_i}(\alpha) + \ln k .$$

(7) **(Continuité de l'entropie des partitions)** Pour toute partition mesurable  $\alpha$  de  $X$  et pour tout  $\eta > 0$ , il existe  $\epsilon > 0$  tel que, si  $\beta$  est une partition mesurable de  $X$  telle que, pour tout  $A$  dans  $\alpha$ , il existe  $B$  dans  $\beta$  avec  $\mu(A \Delta B) \leq \epsilon$ , alors nous avons  $H_\mu(\alpha \mid \beta) \leq \eta$ .

116. Par exemple, si  $(X, \mu)$  est l'intervalle  $[0, 1]$  muni de la mesure de Lebesgue, si  $\alpha = \{[0, \frac{1}{2}[, [\frac{1}{2}, 1]\}$  et  $\beta = \{[0, \frac{1}{4}[, [\frac{1}{4}, 1]\}$ , alors  $H(\alpha) = \ln 2 \simeq 0,693$  et  $H(\beta) = 2 \ln 2 - \frac{3}{4} \ln 3 \simeq 0,462 < H(\alpha)$ .

117. Lorsque  $\beta$  est une partition mesurable dénombrable, il convient de demander que  $H_\mu(\beta) < +\infty$  pour obtenir cette inégalité.

- (8) **(Exhaustion de l'information)** Soit  $(\beta_n)_{n \in \mathbb{N}}$  une suite de partitions mesurables de  $X$  de plus en plus fine (c'est-à-dire vérifiant  $\beta_{n+1} \preceq \beta_n$  pour tout  $n \in \mathbb{N}$ ) et asymptotiquement génératrice (c'est-à-dire telle que la réunion  $\bigcup_{n \in \mathbb{N}} \beta_n$  engendre la  $\sigma$ -algèbre  $\mathcal{B}$ ). Alors  $H_\mu(\alpha | \beta_n)$  tend vers 0 quand  $n \rightarrow \infty$ .

**Démonstration.** (1) Posons  $p = \text{Card } \alpha$ . Par la concavité de  $\varphi$ , nous avons

$$\frac{1}{p} H(\alpha) = \frac{1}{p} \sum_{A \in \alpha} \varphi(\mu(A)) \leq \varphi\left(\frac{1}{p} \sum_{A \in \alpha} \mu(A)\right) = \varphi\left(\frac{1}{p}\right) = \frac{1}{p} \ln p.$$

Le cas d'égalité découle de la concavité stricte de  $\varphi$ .

- (2) La première assertion est immédiate par définition, et la seconde en découle en prenant  $\beta' = \{X\}$ .

Avant de continuer la démonstration de la proposition 5.1, introduisons quelques définitions qui seront utiles pour les démonstrations, mais qu'il n'est pas forcément utile de mémoriser.

L'*information* d'une partition mesurable  $\alpha$  est la fonction  $I_\alpha : X \rightarrow [0, +\infty]$  donnée par

$$I_\alpha(x) = \sum_{A \in \alpha} (-\ln \mu(A)) \mathbb{1}_A(x).$$

L'*information relative* d'une partition mesurable  $\alpha$  par rapport à une partition mesurable  $\beta$  est la fonction

$$I_{\alpha|\beta}(x) = \sum_{A \in \alpha, B \in \beta} -\ln\left(\frac{\mu(A \cap B)}{\mu(B)}\right) \mathbb{1}_{A \cap B}(x).$$

Notons que  $I_{\alpha|\beta} = 0$  si et seulement si  $\alpha \preceq \beta \pmod{0}$  et que si  $\alpha \perp \beta$  alors  $I_{\alpha|\beta} = I_\alpha$ .

L'entropie d'une partition est alors son information moyenne :

$$H_\mu(\alpha) = \int_X I_\alpha(x) d\mu(x) \quad \text{et} \quad H(\alpha | \beta) = \int_X I_{\alpha|\beta}(x) d\mu(x)$$

- (3) Pour tous les  $A \in \alpha$ ,  $B \in \beta$  et  $C \in \gamma$  tels que  $\mu(B \cap C) \neq 0$  et  $\mu(C) \neq 0$ , nous avons

$$\ln\left(\frac{\mu(A \cap B \cap C)}{\mu(C)}\right) = \ln\left(\frac{\mu(A \cap B \cap C)}{\mu(B \cap C)}\right) + \ln\left(\frac{\mu(B \cap C)}{\mu(C)}\right).$$

En tout point de  $A \cap B \cap C$ , nous avons donc  $I_{\alpha \vee \beta | \gamma} = I_{\alpha | \beta \vee \gamma} + I_{\beta | \gamma}$ . La première assertion s'en déduit par intégration, et la seconde en découle en prenant  $\gamma = \{X\}$ .

- (4) (i) Pour tous les  $A \in \alpha$ ,  $B \in \beta$  et  $C \in \gamma$  tel que  $\mu(C) \neq 0$ , si  $B$  est contenu dans  $A$ , alors

$$-\ln\left(\frac{\mu(A \cap C)}{\mu(C)}\right) \leq -\ln\left(\frac{\mu(B \cap C)}{\mu(C)}\right),$$

avec égalité si et seulement si  $\mu((A \setminus B) \cap C) = 0$ . L'inégalité  $I_{\alpha|\gamma} \leq I_{\beta|\gamma}$  s'en déduit. La première assertion du point (i) s'obtient alors par intégration, et la seconde en découle en prenant  $\gamma = \{X\}$ .



(ii) Comme  $\gamma$  est moins fine que  $\beta$ , la partition  $\beta$  induit une partition  $\beta_C$  de tout élément  $C$  de  $\gamma$ . Par la concavité de  $\varphi$ , pour tout  $A \in \alpha$ , nous avons

$$\sum_{B \in \beta_C} \varphi\left(\frac{\mu(A \cap B)}{\mu(B)}\right) \mu(B) \leq \varphi\left(\frac{\mu(A \cap C)}{\mu(C)}\right) \mu(C).$$

Lorsque  $\gamma = \{X\}$ , par la stricte concavité de  $\varphi$ , cette inégalité est une égalité si et seulement si pour tout  $B \in \beta$ , nous avons  $\mu(A \cap B) = \mu(A)\mu(B)$ . En sommant cette inégalité pour  $A$  dans  $\alpha$  et  $C$  dans  $\gamma$ , nous obtenons bien

$$\sum_{A \in \alpha, B \in \beta} \varphi\left(\frac{\mu(A \cap B)}{\mu(B)}\right) \mu(B) \leq \sum_{A \in \alpha, C \in \gamma} \varphi\left(\frac{\mu(A \cap C)}{\mu(C)}\right) \mu(C).$$

La dernière inégalité en découle en prenant  $\gamma = \{X\}$ .

(iii) Ceci découle de la dernière égalité de l'assertion (3) et de la dernière inégalité du point (i), car  $\alpha \preceq \alpha \vee \beta$ .

(5) La première affirmation découle de l'assertion (3) et du point (ii) de l'assertion (4), car  $\gamma \preceq \beta \vee \gamma$ . L'inégalité dans la deuxième affirmation en découle en prenant  $\gamma = \{X\}$ , et son cas d'égalité découle du cas d'égalité dans la question (ii). La dernière affirmation en découle par récurrence.

(6) Montrons que pour toutes les mesures de probabilités  $\nu_1, \dots, \nu_k$  sur  $X'$  et tous les  $s_1, \dots, s_k \in [0, 1]$  tels que  $\sum_{i=1}^k s_i = 1$ , nous avons

$$\sum_{i=1}^k s_i H_{\nu_i}(\alpha) \leq H_{\sum_{i=1}^k s_i \nu_i}(\alpha) \leq \sum_{i=1}^k s_i H_{\nu_i}(\alpha) + \sum_{i=1}^k \varphi(s_i).$$

L'assertion (6) s'en déduit, car par un calcul d'extremum lié, le maximum de  $\sum_{i=1}^k \varphi(s_i)$  sous la contrainte  $\sum_{i=1}^k s_i = 1$  est  $\sum_{i=1}^k \varphi(\frac{1}{k}) = \ln k$ . Montrons la formule centrée ci-dessus pour  $k = 2$ , le cas général s'en déduit par récurrence.

Pour tout  $A$  dans  $\alpha$ , notons  $D_A = \varphi(s_1 \nu_1(A) + s_2 \nu_2(A)) - s_1 \varphi(\nu_1(A)) - s_2 \varphi(\nu_2(A))$ . Par la concavité de  $\varphi$ , nous avons l'inégalité  $D_A \geq 0$ . D'autre part, nous avons

$$\begin{aligned} D_A &= -s_1 \nu_1(A) (\ln(s_1 \nu_1(A) + s_2 \nu_2(A)) - \ln(s_1 \nu_1(A))) \\ &\quad - s_2 \nu_2(A) (\ln(s_1 \nu_1(A) + s_2 \nu_2(A)) - \ln(s_2 \nu_2(A))) \\ &\quad - s_1 \nu_1(A) \ln s_1 - s_2 \nu_2(A) \ln s_2. \end{aligned}$$

Les deux premiers termes après l'égalité sont négatifs, car  $\ln$  est croissante. Donc nous avons  $D_A \leq \nu_1(A) \varphi(s_1) + \nu_2(A) \varphi(s_2)$ . Nous en déduisons que  $0 \leq \sum_{A \in \alpha} D_A \leq \varphi(s_1) + \varphi(s_2)$  comme cherché.

(7) Écrivons  $\alpha = \{A_1, \dots, A_p\}$  où  $p \in \mathbb{N} \setminus \{0\}$ . Les indices  $i$  et  $j$  ci-dessous varient entre 1 et  $p$ . Nous pouvons supposer que  $\mu(A_i)$  est non nul pour tout  $i$ . Soit  $\epsilon \in ]0, \frac{1}{e}]$  tel que  $\epsilon \leq \frac{1}{2} \min_{1 \leq i \leq p} \mu(A_i)$ . Soit  $\beta$  une partition mesurable de  $X$  telle que, pour tout  $i$ , il existe  $B_i$  dans  $\beta$  vérifiant  $\mu(A_i \Delta B_i) \leq \frac{1}{p+1} \epsilon^2$ . Nous avons  $\mu(X \setminus \bigcup_{i=1}^p B_i) \leq \frac{p}{p+1} \epsilon^2$ , car

$$X \setminus \bigcup_{i=1}^p B_i = \bigcup_{j=1}^p (A_j \setminus \bigcup_{i=1}^p B_i) \subset \bigcup_{j=1}^p (A_j \setminus B_j) \subset \bigcup_{j=1}^p (A_j \Delta B_j).$$

Quitte à remplacer  $\beta$  par  $\{B_1 \cup (X \setminus \bigcup_{i=1}^p B_i), B_2, \dots, B_p\}$ , qui est une partition moins fine que  $\beta$ , donc augmente l'entropie relative  $H_\mu(\alpha | \beta)$  par l'assertion (4) (ii), nous pouvons supposer que  $\beta = \{B_1, B_2, \dots, B_p\}$  et que  $\mu(A_i \Delta B_i) \leq \frac{1}{p+1} \epsilon^2 + \frac{p}{p+1} \epsilon^2 = \epsilon^2 \leq \epsilon$  pour tout  $i$ . Notons qu'alors par l'hypothèse sur  $\epsilon$ , pour tout  $i$ , nous avons

$$\mu(B_i) \geq \mu(A_i \cap B_i) = \mu(A_i) - \mu(A_i \setminus B_i) \geq \mu(A_i) - \epsilon \geq \frac{1}{2} \mu(A_i),$$

et donc

$$\mu(B_i \setminus A_i) \leq \mu(A_i \Delta B_i) \leq \epsilon^2 \leq \frac{1}{2} \mu(A_i) \epsilon \leq \mu(B_i) \epsilon.$$

Puisque  $B_i = (A_i \cap B_i) \sqcup (B_i \setminus A_i)$  et  $A_i \cap B_j \subset A_i \setminus B_i$  si  $i \neq j$  car  $B_j \subset {}^c B_i$ , il en découle que

$$\forall i, \quad \frac{\mu(A_i \cap B_i)}{\mu(B_i)} \geq 1 - \epsilon \quad \text{et} \quad \forall i \neq j, \quad \frac{\mu(A_i \cap B_j)}{\mu(B_j)} \leq \epsilon.$$

Puisque  $\epsilon \leq \frac{1}{e}$ , l'application  $\varphi$  est croissante sur  $[0, \epsilon]$  et décroissante sur  $[1 - \epsilon, 1]$ . Donc nous avons

$$H_\mu(\alpha | \beta) = \sum_{i,j} \varphi\left(\frac{\mu(A_i \cap B_j)}{\mu(B_j)}\right) \mu(B_j) \leq p^2 \max\{\varphi(\epsilon), \varphi(1 - \epsilon)\}.$$

Comme ce majorant tend vers 0 avec  $\epsilon$ , le résultat en découle.

(8) Nous commençons par montrer le lemme suivant.

**Lemme 5.2.** *Soient  $(X, \mathcal{B}, \mu)$  un espace de probabilité et  $\mathcal{C}$  une sous-algèbre de Boole de  $\mathcal{B}$  qui engendre la  $\sigma$ -algèbre  $\mathcal{B}$ . Alors, pour tout  $B$  dans  $\mathcal{B}$  et  $\epsilon > 0$ , il existe  $C$  dans  $\mathcal{C}$  telle que  $\mu(B \Delta C) \leq \epsilon$ .*

**Démonstration.** Soit  $\mathcal{B}' = \{B \in \mathcal{B} : \forall \epsilon > 0, \exists C \in \mathcal{C}, \mu(B \Delta C) \leq \epsilon\}$ . Comme  $\mathcal{B}'$  contient  $\mathcal{C}$ , il suffit de montrer que  $\mathcal{B}'$  est une  $\sigma$ -algèbre. Or  $\mathcal{B}'$  est stable par

- passage au complémentaire, car  ${}^c B \Delta {}^c C = B \Delta C$ ,
- union finie, car  $(A_1 \cup A_2) \Delta (C_1 \cup C_2) \subset (A_1 \Delta C_1) \cup (A_2 \Delta C_2)$ ,
- union dénombrable croissante, car si  $B$  est la réunion d'une suite croissante  $(B_n)_{n \in \mathbb{N}}$ , alors  $\mu(B \Delta B_n)$  tend vers 0.  $\square$

Retournons à la démonstration de l'assertion (8). Écrivons  $\alpha = \{A_1, \dots, A_p\}$ . Nous pouvons supposer que  $\mu(A_i)$  est non nul pour tout  $i$ . Nous appliquons le lemme précédent avec  $\mathcal{C}$  l'ensemble des unions finies d'éléments des  $\beta_n$ , qui est une algèbre de Boole, car les  $\beta_n$  sont des partitions de plus en plus fine. Pour tout  $\epsilon > 0$ , nous pouvons trouver  $n_0 \geq 1$  et des unions (finies)  $B_1, \dots, B_p$  d'éléments de  $\beta_{n_0}$  telles que  $\mu(A_i \Delta B_i) \leq \epsilon$ . Comme les  $A_i$  sont disjoints, nous avons  $\mu(B_i \Delta B_j) \leq 2\epsilon$  si  $1 \leq i \neq j \leq p$ . Nous avons

$$\mu\left(X \setminus \bigcup_{i=1}^p B_i\right) = \mu\left(\bigcup_{i=1}^p A_i \setminus \bigcup_{i=1}^p B_i\right) \leq \mu\left(\bigcup_{i=1}^p (A_i \setminus B_i)\right) \leq p\epsilon.$$

Quitte à remplacer  $B_1, B_2, \dots, B_p$  par  $B_1 \cup (X \setminus \bigcup_{i=1}^p B_i), B_2 \setminus B_1, \dots, B_p \setminus \bigcup_{i=1}^{p-1} B_i$  respectivement, qui sont encore des unions finies d'éléments de  $\beta_{n_0}$ , et à remplacer  $\epsilon$  par  $2(p+1)\epsilon$ , nous pouvons supposer que  $\beta = \{B_1, \dots, B_p\}$  est une partition mesurable. Pour

tout  $\eta > 0$ , si  $\epsilon$  est assez petit, par l'assertion (7), nous avons  $H_\mu(\alpha \mid \beta) \leq \eta$ . Par la monotonie de l'entropie des partitions (voir l'assertion (4) (ii)), pour tout  $n \geq n_0$ , comme  $\beta \preceq \beta_{n_0} \preceq \beta_n$ , nous avons

$$0 \leq H_\mu(\alpha \mid \beta_n) \leq H_\mu(\alpha \mid \beta_{n_0}) \leq H_\mu(\alpha \mid \beta) \leq \eta .$$

Ceci montre le résultat.  $\square$

**Remarque.** Notons que l'assertion (8) dit que l'information relative  $I_{\alpha \mid \beta_n}$  tends vers 0 dans  $\mathbb{L}^1(X, \mu)$  et que le théorème des martingales (voir par exemple [LeG]) permet de montrer que  $I_{\alpha \mid \beta_n}$  tends vers 0 presque sûrement pour la mesure  $\mu$  quand  $n \rightarrow +\infty$ .

## 5.2 Entropie d'un système dynamique mesuré

Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret (voir la remarque 5.5 pour le cas du temps continu). Soient  $\alpha$  et  $\beta$  des partitions mesurables de  $X$ .

L'étude de l'entropie des partitions jointes  $\bigvee_{i=0}^n \phi^{-i}\alpha$ , qui permettent de pister les orbites de  $\phi$  jusqu'au temps  $n$ , par la suite des parties de  $\alpha$  qu'elles visitent entre le temps 0 et le temps  $n$ , s'avère fort utile.

Nous appelons *entropie de  $\phi$  relativement à la partition mesurable  $\alpha$*  la limite (qui existe, voir la proposition 5.3 (1) ci-dessous)

$$h_\mu(\phi, \alpha) = \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} \phi^{-i}\alpha\right) . \quad (48)$$

Nous montrerons dans la proposition 5.3 (3) que si  $\alpha \preceq \beta$ , alors  $h_\mu(\phi, \alpha) \leq h_\mu(\phi, \beta)$ . Donc passer à une partition plus fine ne peut que faire croître l'entropie relative à la partition. Ceci motive le passage à la borne supérieure dans la définition suivante.

Nous appelons *entropie mesurée*<sup>118</sup> (ou parfois juste *entropie*) du système dynamique mesuré  $(X, \mathcal{B}, \mu, \phi)$  (ou parfois juste de  $\phi$  lorsque  $(X, \mathcal{B}, \mu)$  est sous-entendu ou parfois juste de  $\mu$  lorsque  $(X, \mathcal{B}, \phi)$  est sous-entendu) la quantité

$$h_\mu(\phi) = \sup_{\alpha} h_\mu(\phi, \alpha) \in [0, +\infty] ,$$

où la borne supérieure est prise sur toutes les partitions mesurables  $\alpha$  de  $(X, \mathcal{B}, \mu)$ .<sup>119</sup>

Une partition mesurable  $\alpha$  de  $X$  est dite *génératrice* pour  $\phi$ , lorsque  $\phi$  n'est pas inversible, si la suite des partitions  $(\alpha_n = \bigvee_{i=0}^n \phi^{-i}\alpha)_{n \in \mathbb{N}}$  est asymptotiquement génératrice (c'est-à-dire si  $\bigcup_{n \in \mathbb{N}} \alpha_n$  engendre la  $\sigma$ -algèbre  $\mathcal{B}$ ), et lorsque  $\phi$  est inversible, si la suite des partitions  $(\alpha_n = \bigvee_{i=-n}^n \phi^{-i}\alpha)_{n \in \mathbb{N}}$  est asymptotiquement génératrice. Voir l'exercice E.33 pour la raison de cette différence.

Nous résumons dans la proposition suivante les principales propriétés de l'entropie mesurée.

**Proposition 5.3.** *Nous avons les propriétés suivantes, pour toutes les partitions mesurables  $\alpha$  et  $\beta$  de  $X$ .*

118. Une terminologie plus classique, mais moins appropriée car il n'y a pas de distance impliquée, est celle d'*entropie métrique*.

119. Lorsque les partitions mesurables dénombrables sont autorisées, il convient de prendre afin de définir  $h_\mu(\phi)$  la borne supérieure sur partitions mesurables  $\alpha$  de  $(X, \mathcal{B}, \mu)$  d'entropie  $H_\mu(\alpha)$  finie.

- (1) **(Existence de l'entropie relative)** La limite dans la formule (48) existe, elle appartient à  $[0, +\infty[$  et nous avons de plus  $h_\mu(\phi, \alpha) = \inf_{n \in \mathbb{N} - \{0\}} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} \phi^{-i} \alpha)$ . Si  $\phi$  est inversible, alors nous avons aussi

$$h_\mu(\phi, \alpha) = \lim_{n \rightarrow +\infty} \frac{1}{2n+1} H_\mu\left(\bigvee_{i=-n}^n \phi^{-i} \alpha\right) = \inf_{n \in \mathbb{N}} \frac{1}{2n+1} H_\mu\left(\bigvee_{i=-n}^n \phi^{-i} \alpha\right).$$

- (2) **(Sous-additivité de l'entropie relative)** Nous avons

$$h_\mu(\phi, \alpha \vee \beta) \leq h_\mu(\phi, \alpha) + h_\mu(\phi, \beta).$$

- (3) **(Monotonie de l'entropie relative)** Si  $\alpha \preceq \beta$ , alors  $h_\mu(\phi, \alpha) \leq h_\mu(\phi, \beta)$ .

- (4) **(Inégalité de Rokhlin <sup>120</sup>)** Nous avons

$$h_\mu(\phi, \alpha) - h_\mu(\phi, \beta) \leq H_\mu(\alpha | \beta).$$

- (5) **(Invariance de l'entropie par conjugaison mesurée)** Si deux systèmes dynamiques probabilisés à temps discret  $(X, \mathcal{B}, \mu, \phi)$  et  $(X', \mathcal{B}', \mu', \phi')$  sont conjugués, alors

$$h_\mu(\phi) = h_{\mu'}(\phi').$$

- (6) **(Multiplicativité de l'entropie)** Pour tout  $k \in \mathbb{N}$ , nous avons  $h_\mu(\phi^k) = k h_\mu(\phi)$ , et si  $\phi$  est inversible, alors  $h_\mu(\phi^k) = |k| h_\mu(\phi)$  pour tout  $k \in \mathbb{Z}$ .

- (7) **(Dépendance affine de l'entropie en la mesure)** Pour tous les  $t \in [0, 1]$  et toutes les mesures de probabilité  $\mu, \nu$  invariantes par  $\phi$  sur un espace mesurable  $(X, \mathcal{B})$ , nous avons  $h_{t\mu+(1-t)\nu}(\phi, \alpha) = t h_\mu(\phi, \alpha) + (1-t) h_\nu(\phi, \alpha)$  et

$$h_{t\mu+(1-t)\nu}(\phi) = t h_\mu(\phi) + (1-t) h_\nu(\phi).$$

- (8) **(Théorème de Kolmogorov-Sinai)** Si une partition mesurable  $\alpha$  est génératrice pour  $\phi$ , alors

$$h_\mu(\phi) = h_\mu(\phi, \alpha).$$

L'inégalité de Rokhlin est un outil très utile dans les estimations d'entropies relatives à des partitions mesurables, voir par exemple l'exercice E.35.

Le théorème de Kolmogorov-Sinai est l'un des moyens les plus pratiques de calculer des entropies mesurées, voir par exemple [PY, Wal, CorFS, Pet, BenP, EW, Cou], ainsi que les exercices E.71 et E.72.

**Démonstration.** (1) Pour montrer le premier résultat, nous allons appliquer le lemme de Fekete 2.7 à la suite  $(a_n = H_\mu(\bigvee_{i=0}^{n-1} \phi^{-i} \alpha))_{n \geq 1}$ . En effet, par la propriété de sous-additivité que nous rappelons

$$\forall \alpha', \beta' \in \mathcal{P}, \quad H_\mu(\alpha' \vee \beta') \leq H_\mu(\alpha') + H_\mu(\beta')$$



120. Vladimir Rokhlin  
1919-1984



Donald Ornstein  
1934-



David Ruelle  
1935-



Mark Pinsker  
1925-2003

de l'entropie des partitions (voir la proposition 5.1 (5)), et par l'invariance de l'entropie des partitions par la transformation  $\phi$  qui préserve la mesure (voir la proposition 5.1 (2)), nous avons

$$a_{m+n} \leq H_\mu\left(\bigvee_{i=0}^{m-1} \phi^{-i}\alpha\right) + H_\mu\left(\bigvee_{i=m}^{m+n-1} \phi^{-i}\alpha\right) = a_m + a_n .$$

La démonstration est analogue lorsque  $\phi$  est inversible.

(2) Par la formule (45) et l'associativité de  $\vee$ , ceci découle de la sous-additivité de l'entropie des partitions (voir la proposition 5.1 (5)) par passage à la limite.

(3) Par les formules (45) et (46), ceci découle de la monotonie de l'entropie des partitions (voir la proposition 5.1 (4) (i)).

(4) Par les assertions (4) (iii), (5), (4) (ii) et (2) (avec l'invariance de  $\mu$  par  $\phi$ ) de la proposition 5.1, nous avons

$$\begin{aligned} H_\mu\left(\bigvee_{i=0}^{n-1} \phi^{-i}\alpha\right) - H_\mu\left(\bigvee_{j=0}^{n-1} \phi^{-j}\beta\right) &\leq H_\mu\left(\bigvee_{i=0}^{n-1} \phi^{-i}\alpha \mid \bigvee_{j=0}^{n-1} \phi^{-j}\beta\right) \leq \sum_{i=0}^{n-1} H_\mu(\phi^{-i}\alpha \mid \bigvee_{j=0}^{n-1} \phi^{-j}\beta) \\ &\leq \sum_{i=0}^{n-1} H_\mu(\phi^{-i}\alpha \mid \phi^{-i}\beta) = nH_\mu(\alpha \mid \beta) . \end{aligned}$$

D'où le résultat en divisant par  $n$  et en passant à la limite.

(5) Soit  $\theta : X \rightarrow X'$  une bijection (quitte à enlever des ensembles de mesure nulle) mesurable et préservant la mesure, ainsi que son inverse, telle que  $\theta \circ \phi = \phi' \circ \theta$ . Si  $\alpha'$  est une partition mesurable de  $X'$ , alors  $\theta^{-1}\alpha' = \{\theta^{-1}(A) : A \in \alpha'\}$  est une partition mesurable de  $X$ . La naturalité de l'entropie des partitions  $H_\mu$  et le fait que  $\theta_*\mu = \mu'$ , puis la propriété de conjugaison de  $\theta$ , impliquent que

$$H_{\mu'}\left(\bigvee_{i=0}^{n-1} (\phi')^{-i}\alpha'\right) = H_\mu(\theta^{-1}\bigvee_{i=0}^{n-1} (\phi')^{-i}\alpha') = H_\mu\left(\bigvee_{i=0}^{n-1} \phi^{-i}\theta^{-1}\alpha'\right) .$$

En divisant par  $n$  et en prenant la limite quand  $n$  tend  $+\infty$ , nous en déduisons que  $h_{\mu'}(\phi', \alpha') = h_\mu(\phi, \theta^{-1}\alpha')$ . En prenant la borne supérieure sur  $\alpha'$ , nous obtenons donc que  $h_{\mu'}(\phi') \leq h_\mu(\phi_\mu)$ . Par symétrie, l'assertion (5) en résulte.

(6) En utilisant l'assertion (3) pour la première inégalité, nous avons

$$h_\mu(\phi^k, \alpha) \leq h_\mu(\phi^k, \bigvee_{i=0}^{k-1} \phi^{-i}\alpha) = \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{nk-1} \phi^{-i}\alpha\right) = k h_\mu(\phi, \alpha) .$$

Ceci prouve que  $h_\mu(\phi^k, \alpha) \leq k h_\mu(\phi, \alpha)$ , mais aussi que

$$k h_\mu(\phi, \alpha) = h_\mu(\phi^k, \bigvee_{i=0}^{k-1} \phi^{-i}\alpha) \leq h_\mu(\phi^k) .$$

par la définition de  $h_\mu(\cdot)$  comme borne supérieure. Donc en prenant la borne supérieure sur  $\alpha$ , nous avons  $h_\mu(\phi^k) = k h_\mu(\phi)$ . Le résultat dans le cas inversible découle de l'exercice E.31.

(7) Appliquons l'assertion (6) de la proposition 5.1 avec  $k = 2$  à la partition mesurable  $\bigvee_{i=0}^{n-1} \phi^{-i} \alpha$ , divisons par  $n$  et prenons la limite quand  $n \rightarrow +\infty$ . La première affirmation en découle car  $\lim_{n \rightarrow +\infty} \frac{\ln(2)}{n} = 0$ .

Pour montrer la deuxième affirmation, pour tout  $n \in \mathbb{N}$ , notons  $(\alpha_n)_{n \in \mathbb{N}}, (\beta_n)_{n \in \mathbb{N}}, (\gamma_n)_{n \in \mathbb{N}}$  des suites de partitions mesurables de  $X$  telles que

$$\lim_{n \rightarrow \infty} h_\mu(\phi, \alpha_n) = h_\mu(\phi), \quad \lim_{n \rightarrow \infty} h_\nu(\phi, \beta_n) = h_\nu(\phi), \quad \lim_{n \rightarrow \infty} h_{s\mu+t\nu}(\phi, \gamma_n) = h_{s\mu+t\nu}(\phi).$$

Nous pouvons, sans changer ces égalités, remplacer ces partitions par des partitions plus fines comme  $\alpha_n \vee \beta_n \vee \gamma_n$ . Nous pouvons donc supposer que  $\alpha_n = \beta_n = \gamma_n$ . La deuxième affirmation se déduit alors par passage à la limite de la première.

(8) Nous commençons par montrer le lemme suivant. Pour tout  $n \in \mathbb{N}$ , posons  $\alpha_n = \bigvee_{i=0}^n \phi^{-i} \alpha$  et aussi  $\alpha_n = \bigvee_{i=-n}^n \phi^{-i} \alpha$  dans le cas inversible.

**Lemme 5.4.** *Pour tout  $k \in \mathbb{N}$ , nous avons  $h_\mu(\phi, \alpha) = h_\mu(\phi, \bigvee_{i=0}^k \phi^{-i} \alpha)$ . De plus, si  $\phi$  est inversible, alors  $h_\mu(\phi, \alpha) = h_\mu(\phi, \bigvee_{i=-k}^k \phi^{-i} \alpha)$ .*

**Démonstration.** Supposons tout d'abord que  $\alpha_n = \bigvee_{i=0}^n \phi^{-i} \alpha$  pour tout  $n \in \mathbb{N}$ . Notons  $a_n(\alpha) = H_\mu(\alpha_{n-1})$ . Comme

$$\bigvee_{i=0}^{n-1} \phi^{-i} \alpha_k = \bigvee_{i=0}^{n+k-1} \phi^{-i} \alpha,$$

nous avons  $a_n(\alpha_k) = a_{n+k}(\alpha)$ . Donc les suites  $\frac{1}{n} a_n(\alpha_k)$  et  $\frac{1}{n} a_n(\alpha)$  ont la même limite. Nous procédons de même dans le cas inversible.  $\square$

Reprenons la démonstration de l'assertion (8). Fixons  $\epsilon > 0$ . Par la définition de  $h_\mu(\phi)$  comme borne supérieure, il existe une partition mesurable  $\beta$  telle que  $h_\mu(\phi, \beta) \geq h_\mu(\phi) - \epsilon$ . Puisque la partition mesurable  $\alpha$  est génératrice pour  $\phi$ , d'après la proposition 5.1 (8), il existe  $n \geq 1$  tel que  $H_\mu(\beta | \alpha_n) \leq \epsilon$ . Mais alors, par le lemme 5.4 ci-dessus puis l'assertion (4) ci-dessus, nous avons

$$h_\mu(\phi) \geq h_\mu(\phi, \alpha) = h(\phi, \alpha_n) \geq h_\mu(\phi, \beta) - H_\mu(\beta | \alpha_n) \geq h_\mu(\phi) - 2\epsilon.$$

Ceci est vrai pour tout  $\epsilon > 0$ , donc  $h_\mu(\phi) = h_\mu(\phi, \alpha)$ .  $\square$

**Remarque 5.5.** Si  $(X, \mathcal{B}, \mu, \phi = (\phi^t)_{t \in \mathbb{R}})$  est un système dynamique probabilisé à temps continu, alors pour tout  $t \in \mathbb{R}$  fixé, nous pouvons considérer le système dynamique probabilisé à temps discret  $(X, \mathcal{B}, \mu, \phi^t)$ , et calculer l'entropie  $h_\mu(\phi^t, \alpha)$  relative à une partition mesurable  $\alpha$  et l'entropie mesurée  $h_\mu(\phi^t)$ . Comme dans la démonstration de l'assertion 6 ci-dessus, nous avons

$$\forall t \in \mathbb{R}, \quad h_\mu(\phi^t) = |t| h_\mu(\phi^1).$$

Pour cette raison, l'entropie mesurée d'un système dynamique probabilisé à temps continu est souvent définie comme l'entropie mesurée de son temps 1.

**Exemple :** Soient  $\mathcal{A}$  un alphabet fini,  $\nu$  une mesure de probabilité sur  $\mathcal{A}$ ,  $X$  l'espace borélien produit  $\mathcal{A}^{\mathbb{N}}$  ou  $\mathcal{A}^{\mathbb{Z}}$ ,  $\mu$  la mesure de probabilité produit  $\nu^{\mathbb{N}}$  ou  $\nu^{\mathbb{Z}}$  sur  $X$  et  $\phi$  le décalage à gauche  $\sigma_+$  ou  $\sigma$  sur  $X$ . Posons  $p_a = \nu(\{a\})$  pour tout  $a$  dans  $\mathcal{A}$ . Alors l'entropie

mesurée du système de Bernoulli  $(X, \mu, \phi)$  muni de sa mesure de Bernoulli associé à  $(\mathcal{A}, \nu)$  est

$$h_{\nu^{\mathbb{N}}}(\sigma_+) = h_{\nu^{\mathbb{Z}}}(\sigma) = - \sum_{a \in \mathcal{A}} p_a \ln p_a . \quad (49)$$

Montrons ce résultat pour le cas bilatère (inversible), le cas unilatère est analogue. Pour tout  $a$  dans  $\mathcal{A}$ , posons  $[a] = \{(x_i)_{i \in \mathbb{Z}} \in X : x_0 = a\}$ . Soit  $\alpha = \{[a] : a \in \mathcal{A}\}$  la partition mesurable par les cylindres centrés de longueur 1. Alors  $\alpha$  est génératrice, car pour tout  $n \in \mathbb{N}$ , la partition mesurable  $\bigvee_{i=-n}^n \sigma^{-i} \alpha$  est la partition par les cylindres centrés de longueur  $2n + 1$

$$\bigvee_{i=-n}^n \sigma^{-i} \alpha = \{[a_{-n}, \dots, a_n] = \{(x_i)_{i \in \mathbb{Z}} \in X : x_{-n} = a_{-n}, \dots, x_n = a_n\} : a_{-n}, \dots, a_n \in \mathcal{A}\}$$

et puisque la  $\sigma$ -algèbre de  $X$  est engendrée par l'ensemble de tous ces cylindres. Nous calculons alors facilement, en développant le logarithme et en utilisant que  $\sum_{a_k \in \mathcal{A}} p_{a_k} = 1$  pour tout  $k = -n \dots, n$  que

$$\begin{aligned} H_{\nu^{\mathbb{Z}}}\left(\bigvee_{i=-n}^n \sigma^{-i} \alpha\right) &= - \sum_{a_{-n}, \dots, a_n \in \mathcal{A}} \mu([a_{-n}, \dots, a_n]) \ln \mu([a_{-n}, \dots, a_n]) \\ &= - \sum_{a_{-n}, \dots, a_n \in \mathcal{A}} (p_{a_{-n}} \dots p_{a_n}) \ln(p_{a_{-n}} \dots p_{a_n}) \\ &= -(2n + 1) \sum_{a \in \mathcal{A}} p_a \ln p_a . \end{aligned}$$

Le résultat s'en déduit.

En particulier, si  $\nu$  est la loi uniforme sur  $\mathcal{A}$  et si  $p$  est le cardinal de  $\mathcal{A}$ , alors

$$h_{\nu^{\mathbb{N}}}(\sigma_+) = h_{\nu^{\mathbb{Z}}}(\sigma) = \log p .$$

Par l'invariance de l'entropie par conjugaison (voir l'assertion (5) de la proposition 5.3), deux systèmes de Bernoulli munis de leur mesure de Bernoulli sur des alphabets finis non vide de cardinaux différents munis de leur loi uniforme ne sont pas conjugués (en tant que systèmes dynamiques probabilisés). Nous reviendrons sur cet exemple (pour le généraliser) au paragraphe 6.1. Pour d'autres calculs, voir les exercices ci-dessous.

L'entropie des systèmes de Bernoulli munis de leur mesure de Bernoulli possède un intérêt historique très important, car leur étude a été un moteur important du développement de la théorie ergodique. Les résultats les plus frappants sont les suivants, dont nous n'aurons pas le temps de donner un aperçu de démonstration dans ces notes. Nous renvoyons par exemple à [DoS] pour une démonstration de ces deux résultats, ainsi que pour une présentation historique.

**Théorème 5.6. (Théorème du facteur de Sinai)** *Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret ergodique. Tout système de Bernoulli bilatère ou unilatère muni de sa mesure de Bernoulli, sur un alphabet fini muni d'une mesure de probabilité, d'entropie mesurée inférieure ou égale à  $h_{\mu}(\phi)$  est un facteur de  $(X, \mathcal{B}, \mu, \phi)$ .*

**(Théorème d'isomorphisme d'Ornstein<sup>120</sup>)** *Deux systèmes de Bernoulli bilatères munis de leur mesure de Bernoulli  $(\mathcal{A}_1^{\mathbb{Z}}, \nu_1^{\mathbb{Z}}, \sigma)$  et  $(\mathcal{A}_2^{\mathbb{Z}}, \nu_2^{\mathbb{Z}}, \sigma)$ , sur des alphabets finis muni d'une mesure de probabilité  $(\mathcal{A}_1, \nu_1)$  et  $(\mathcal{A}_2, \nu_2)$  respectivement, sont conjugués si et seulement si leurs entropies mesurées sont égales.  $\square$*

La condition d'entropie inférieure dans le théorème du facteur de Sinai est nécessaire par l'exercice E.32 (1). La nécessité de l'égalité des entropies pour être conjugués est un fait général qui a été vu dans la proposition 5.3 (5). Le caractère remarquable du théorème d'Orstein est qu'il dit que l'entropie est un invariant complet (permettant la classification à conjugaison près) sur la famille des systèmes de Bernoulli. Mais il s'est avéré que l'entropie était loin d'être un invariant complet sur toute famille de systèmes dynamiques probabilisés à temps discret ergodiques.

Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret. Rappelons que  $\mathcal{P}_{\text{mod}_\mu 0}$  est l'ensemble quotient de l'ensemble  $\mathcal{P}$  des partitions mesurables de  $(X, \mathcal{B})$  par la relation d'équivalence « être égales  $\mu$ -presque partout ». <sup>121</sup> Pour toutes les  $\alpha, \beta \in \mathcal{P}$ , posons

$$d(\alpha, \beta) = H_\mu(\alpha | \beta) + H_\mu(\beta | \alpha).$$

**Proposition 5.7.** *La quantité  $d(\alpha, \beta)$  ne dépend que de la classe d'équivalence de  $\alpha, \beta \in \mathcal{P}$  pour l'égalité  $\mu$ -presque partout. L'application induite  $d : \mathcal{P}_{\text{mod}_\mu 0} \times \mathcal{P}_{\text{mod}_\mu 0} \rightarrow [0, +\infty[$  est une distance sur  $\mathcal{P}_{\text{mod}_\mu 0}$ , appelée la distance de Rokhlin.*

(1) *Pour toutes les  $\alpha, \beta \in \mathcal{P}$ , nous avons  $|H_\mu(\alpha) - H_\mu(\beta)| \leq d(\alpha, \beta)$  et*

$$|h_\mu(\phi, \alpha) - h_\mu(\phi, \beta)| \leq d(\alpha, \beta).$$

(2) *Pour tous les  $N \in \mathbb{N}$  et  $\epsilon > 0$ , il existe  $\eta > 0$  tel que pour toutes les partitions mesurables  $\alpha, \beta \in \mathcal{P}$  de même cardinal  $N$ , si pour tout  $A \in \alpha$ , il existe  $B \in \beta$  tel que  $\mu(A \Delta B) \leq \eta$  et pour tout  $B \in \beta$ , il existe  $A \in \alpha$  tel que  $\mu(A \Delta B) \leq \eta$ , alors  $d(\alpha, \beta) \leq \epsilon$ .*

(3) *Soit  $\Xi$  une partie de  $\mathcal{P}$  telle que l'ensemble des partitions mesurables moins fine qu'un élément de  $\Xi$  soit d'image dense dans l'espace quotient  $\mathcal{P}_{\text{mod}_\mu 0}$  muni de la distance de Rokhlin. Alors*

$$h_\mu(\phi) = \sup_{\xi \in \Xi} h_\mu(\phi, \xi).$$

La dernière affirmation (3) est un renforcement du théorème de Kolmogorov-Sinai, car si  $\alpha$  est une partition mesurable génératrice, alors l'ensemble  $\Xi = \{\bigvee_{i=0}^n \phi^{-i} \alpha : n \in \mathbb{N}\}$  vérifie les hypothèses de cette dernière affirmation. Par exemple, il découle des propriétés de régularité des mesures finies sur un espace localement compact, et des assertions (2) et (3) ci-dessus, que si  $X$  est un espace métrique localement compact, et si  $\Xi = \{\alpha_n : n \in \mathbb{N}\}$  est telle que  $\lim_{n \rightarrow +\infty} \sup_{A \in \alpha_n} \text{diam}(A) = 0$ , alors  $h_\mu(\phi) = \lim_{n \rightarrow +\infty} h_\mu(\phi, \alpha_n)$ .

**Démonstration.** L'application  $d : \mathcal{P} \times \mathcal{P} \rightarrow [0, +\infty[$  est positive ou nulle, et symétrique. Par l'invariance de l'entropie relative des partitions par égalité presque partout (voir la formule (47)), elle passe au quotient en une application  $d : \mathcal{P}_{\text{mod}_\mu 0} \times \mathcal{P}_{\text{mod}_\mu 0} \rightarrow [0, +\infty[$ , qui est positive ou nulle, et symétrique.

Pour toutes les  $\alpha, \beta, \gamma \in \mathcal{P}$ , montrons que

$$H_\mu(\alpha | \gamma) \leq H_\mu(\alpha | \beta) + H_\mu(\beta | \gamma). \quad (50)$$

121. Lorsque les partitions mesurables dénombrables sont autorisées, il convient de ne prendre, afin de définir  $\mathcal{P}_{\text{mod}_\mu 0}$ , que les partitions mesurables  $\alpha$  de  $(X, \mathcal{B}, \mu)$  d'entropie  $H_\mu(\alpha)$  finie.



En effet, en utilisant respectivement, dans la proposition 5.1, l'assertion (3), l'assertion (4) (i) puis (ii) car  $(\alpha \vee \beta) \preceq \beta$ , puis trois fois l'assertion (3), nous avons

$$\begin{aligned} H(\alpha \mid \gamma) &= H(\alpha \vee \gamma) - H(\gamma) \leq H(\alpha \vee \beta \vee \gamma) - H(\gamma) \\ &\leq H(\alpha \vee \beta \vee \gamma) + (-H(\gamma \mid \alpha \vee \beta) + H(\gamma \mid \beta)) - H(\gamma) \\ &= H(\alpha \vee \beta) + H(\gamma \mid \beta) - H(\gamma) = H(\alpha \vee \beta) + H(\gamma \vee \beta) - H(\beta) - H(\gamma) \\ &= H(\alpha \mid \beta) + H(\beta \mid \gamma). \end{aligned}$$

L'inégalité triangulaire découle alors de la formule (50), car

$$\begin{aligned} d(\alpha, \gamma) &= H_\mu(\alpha \mid \gamma) + H_\mu(\gamma \mid \alpha) \leq H_\mu(\alpha \mid \beta) + H_\mu(\beta \mid \gamma) + H_\mu(\gamma \mid \beta) + H_\mu(\beta \mid \alpha) \\ &= d(\alpha, \beta) + d(\beta, \gamma). \end{aligned}$$

(1) Les deux affirmations suivantes dans l'énoncé 5.7 découlent immédiatement par symétrisation de la proposition 5.1 (4) (iii) et de la proposition 5.3 (4).

(2) L'avant-dernière affirmation découle aussi par symétrisation des arguments de la proposition 5.1 (7). Plus précisément, par la continuité de l'application  $\varphi : t \mapsto -t \ln t$ , il existe  $\delta > 0$  tel que si  $t \in [0, \delta] \cup [1 - \delta, 1]$ , alors  $\varphi(t) \leq \frac{\epsilon}{2(N^2+1)}$ . Soient  $\alpha$  et  $\beta$  des partitions mesurables de  $X$ . Pour tout élément  $A \in \alpha$ , il existe au plus un élément  $B \in \beta$  tel que  $\mu(A \Delta B) < \delta$  et  $\mu(A \cap B) \geq \delta$ . Notons  $I_\delta$  l'ensemble des couples  $(A, B) \in \alpha \times \beta$  tels que  $\mu(A \cap B) \leq \delta$ . Si  $\eta < \frac{\delta}{N}$ , et si  $\alpha, \beta$  vérifient l'hypothèse de l'assertion (2), alors  $C = \bigcup_{(A,B) \in (\alpha \times \beta) \setminus I_\delta} A \cap B$  est de mesure au moins  $1 - \delta$ . Soit  $\gamma$  la partition mesurable  $\{A \cap B : (A, B) \in I_\delta\} \cup \{C\}$ , qui contient au plus  $N^2 + 1$  éléments. Alors par les assertions (3) et (4) (ii) de la proposition 5.1, nous avons

$$\begin{aligned} H_\mu(\alpha \mid \beta) &= H_\mu(\alpha \vee \beta) - H_\mu(\beta) = H_\mu(\gamma \vee \beta) - H_\mu(\alpha) = H_\mu(\gamma \mid \beta) \leq H_\mu(\gamma) \\ &\leq (N^2 + 1) \frac{\epsilon}{2(N^2 + 1)} = \frac{\epsilon}{2}. \end{aligned}$$

L'argument étant symétrique, nous avons donc  $d(\alpha, \beta) \leq \epsilon$ .

(3) Pour montrer la dernière affirmation, soit  $\alpha$  une partition mesurable quelconque. Pour tout  $\epsilon > 0$ , par l'hypothèse de (3), il existe une partition mesurable  $\beta$  moins fine qu'un élément  $\xi \in \Xi$  telle que  $d(\alpha, \beta) \leq \epsilon$ . Donc par l'assertion (1) et par la monotonie de l'entropie relative (voir la proposition 5.3 (3)), nous avons

$$h_\mu(\phi, \alpha) \leq h_\mu(\phi, \beta) + \epsilon \leq h_\mu(\phi, \xi) + \epsilon \leq \sup_{\xi \in \Xi} h_\mu(\phi, \xi) + \epsilon.$$

Donc  $h_\mu(\phi, \alpha) \leq \sup_{\xi \in \Xi} h_\mu(\phi, \xi)$  en faisant tendre  $\epsilon$  vers 0 et le résultat s'en déduit en prenant la borne supérieure sur les  $\alpha \in \mathcal{P}$ .  $\square$

### 5.3 L'inégalité de Ruelle <sup>120</sup>

Nous concluons cette partie sur l'entropie mesurée en donnant un critère utilisant les exposants de Lyapounov pour montrer la positivité d'une entropie d'un système dynamique différentiable. Le résultat suivant est valable bien plus généralement (pour toutes les variétés compactes au lieu du tore), mais la majeure partie des idées est présente dans le cas suivant.

Fixons  $N \in \mathbb{N} \setminus \{0\}$ , et considérons le système dynamique différentiable  $(\mathbb{T}^N, \phi)$  où  $\mathbb{T}^N = \mathbb{R}^N / \mathbb{Z}^N$  est le tore de dimension  $N$  et  $\phi : \mathbb{T}^N \rightarrow \mathbb{T}^N$  un  $C^1$ -difféomorphisme. Nous noterons  $\tilde{\phi} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  un  $C^1$ -difféomorphisme  $\mathbb{Z}^N$ -périodique relevant  $\phi$  par la projection canonique  $\uparrow : \mathbb{R}^N \rightarrow \mathbb{T}^N$ , c'est-à-dire tel que le diagramme suivant commute :

$$\begin{array}{ccc} \mathbb{R}^N & \xrightarrow{\tilde{\phi}} & \mathbb{R}^N \\ \uparrow \downarrow & & \downarrow \uparrow \\ \mathbb{T}^N & \xrightarrow{\phi} & \mathbb{T}^N \end{array} .$$

Notons  $M : \mathbb{T}^N \rightarrow \text{GL}_N(\mathbb{R})$  le cocycle linéaire continu (donc mesurable) défini par

$$M : x \mapsto d\tilde{\phi}_{\tilde{x}} ,$$

où  $\tilde{x} \in \uparrow(x)$  et où nous identifions comme d'habitude un opérateur linéaire de  $\mathbb{R}^N$  avec sa matrice dans la base canonique de  $\mathbb{R}^N$ . Il ne dépend ni du choix du relevé  $\tilde{x}$  de  $x$  par  $\mathbb{Z}^N$ -périodicité, ni du relevé  $\tilde{\phi}$  de  $\phi$ , car deux relevés de  $\phi$  diffèrent d'une constante additive. Notons  $(M_n)_{n \in \mathbb{N}}$  le produit dynamique de matrice associé à ce cocycle linéaire, de sorte que, par le théorème de dérivation des fonctions composées, nous ayons

$$\forall x \in \mathbb{T}^N, \quad M_n(x) = d(\tilde{\phi}^n)_{\tilde{x}} = d\tilde{\phi}_{\tilde{\phi}^{n-1}(\tilde{x})} \circ \cdots \circ d\tilde{\phi}_{\tilde{\phi}(\tilde{x})} \circ d\tilde{\phi}_{\tilde{x}} .$$

Fixons une mesure de probabilité  $\phi$ -invariante ergodique  $\mu$  sur  $\mathbb{T}^N$ . Le théorème d'Oseledets 2.12, dans le cas inversible et ergodique, donne donc des exposants de Lyapounov constants en  $\mu$ -presque tout points de  $\mathbb{T}^N$ , de valeurs notées  $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ , et pour  $\mu$ -presque tout  $x \in X$ , des sous-espaces vectoriels  $E_1(x), \dots, E_k(x)$  en somme directe dans  $\mathbb{R}^N$ , de dimensions constantes notées  $\dim E_1, \dots, \dim E_k$ , tels que pour une norme  $\| \cdot \|$  indifférente sur  $\mathbb{R}^N$  et pour tout  $v \in E_i(x) - \{0\}$ , nous ayons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|M_n(x)v\| = \lambda_i .$$

et, en notant  $\sigma_1(A), \dots, \sigma_N(A)$  les valeurs singulières (ordonnées en décroissant, voir la partie 2.4 pour des rappels) d'une matrice  $A \in \text{GL}_N(\mathbb{R})$ , pour tous les entiers  $i \in \llbracket 1, N \rrbracket$  et  $j \in \llbracket 1 + \sum_{\ell=1}^{i-1} \dim E_\ell, \sum_{\ell=1}^i \dim E_\ell \rrbracket$ , nous ayons

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \ln \sigma_j(M_n(x)) = \lambda_i .$$

**Théorème 5.8. (Inégalité de Ruelle)** *Si  $\phi$  est de classe  $C^2$ , nous avons*

$$h_\mu(\phi) \leq \sum_{i=1}^k (\dim E_i) \max\{0, \lambda_i\} .$$

Munissons  $\mathbb{R}^N$  de sa norme (donc de sa distance  $d$ ) euclidienne usuelle et  $\mathcal{M}_N(\mathbb{R})$  de la norme d'opérateur associée. Comme  $\sigma_j(M_n(x)) \leq \|M_n(x)\| \leq \sup_{\tilde{x} \in \mathbb{R}^N} \|d\tilde{\phi}_{\tilde{x}}\|^n$ , il découle de l'inégalité de Ruelle que

$$h_\mu(\phi) \leq N \ln^+ \left( \sup_{\tilde{x} \in \mathbb{R}^N} \|d\tilde{\phi}_{\tilde{x}}\| \right) .$$

**Démonstration.** Nous munissons  $\mathbb{T}^N$  de la distance quotient  $d$  définie par

$$(x, y) \mapsto \min_{\tilde{x} \in \mathcal{P}^{-1}(x), \tilde{y} \in \mathcal{P}^{-1}(y)} d(\tilde{x}, \tilde{y}),$$

de sorte que l'application  $\mathcal{P} : \mathbb{R}^N \rightarrow \mathbb{T}^N$  soit une isométrie en restriction à toute boule fermée de rayon au plus  $\frac{1}{4}$ . Dans ce qui suit, nous prendrons  $\epsilon > 0$  assez petit pour que les parties considérées de  $\mathbb{R}^N$ , dépendant de  $\epsilon$  et dont le diamètre tend vers 0 quand  $\epsilon \rightarrow 0$ , s'envoient isométriquement sur  $\mathbb{T}^N$  par  $\mathcal{P}$ . En notant, pour tout  $j \in \llbracket 1, N \rrbracket$ ,

$$\widehat{\lambda}_j = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \sigma_j(M_n(x))$$

et  $\widehat{\lambda}_j^+ = \max\{0, \widehat{\lambda}_j\}$ , il s'agit de montrer que

$$h_\mu(\phi) \leq \sum_{j=1}^N \widehat{\lambda}_j^+.$$

Le lemme suivant est un argument classique de majoration du volume d'un ellipsoïde. Nous notons  $\mathcal{V}_\epsilon(P)$  le  $\epsilon$ -voisinage fermé d'une partie  $P$  de  $\mathbb{R}^N$ .

**Lemme 5.9.** *Il existe une constante  $c_N > 0$  telle que pour tous les  $A \in \text{GL}_N(\mathbb{R})$  et  $r > 0$ , le nombre maximal  $n(r)$  de boules euclidiennes de rayon  $r/2$  deux à deux disjointes rencontrant  $\mathcal{V}_{3r}(A(B_{\mathbb{R}^N}(0, r)))$  est au plus  $c_N \prod_{i=1}^N \max\{\sigma_i(A), 1\}$ .*

**Démonstration.** Nous pouvons par homogénéité supposer que  $r = 1$  et par invariance de la norme par le groupe orthogonal que  $A$  est diagonale (voir les rappels sur la décomposition en valeurs singulières dans la partie 2.4). Alors  $A(B_{\mathbb{R}^N}(0, 1))$  est l'ellipsoïde d'équation  $\sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} \leq 1$ . Toute boule  $B$  de rayon  $1/2$  est de diamètre 1, donc est contenue dans  $\mathcal{V}_4(A(B_{\mathbb{R}^N}(0, 1)))$  si elle rencontre  $\mathcal{V}_3(A(B_{\mathbb{R}^N}(0, 1)))$ , et son volume  $\text{vol}(B)$  est une constante  $V_N$  qui ne dépend que de  $N$ . Donc  $n(1) \leq \frac{1}{V_N} \text{vol}(\mathcal{V}_4(A(B_{\mathbb{R}^N}(0, 1))))$ , et il suffit par conséquent de majorer le volume du voisinage  $\mathcal{V}_4(A(B_{\mathbb{R}^N}(0, 1)))$ . Or celui-ci est contenu dans le parallélépipède  $\prod_{i=1}^N [-\sigma_i - 4, \sigma_i + 4]$ , de volume  $2^N \prod_{i=1}^N (\sigma_i + 4) \leq 10^N \prod_{i=1}^N (\max\{\sigma_i, 1\})$ .  $\square$

Fixons  $n \in \mathbb{N}$ . Nous allons estimer l'entropie  $h_\mu(\phi^n) = n h_\mu(\phi)$  en fonction des valeurs singulières en construisant des partitions mesurables appropriées du tore. Par compacité et puisque  $\phi$  est de classe  $C^2$ , par uniforme continuité et par le théorème des accroissements finis, pour tout  $\epsilon \in ]0, \frac{1}{4}]$  assez petit (dépendant de  $n$ ), pour tous les  $\tilde{x}, \tilde{y} \in [-1, 2]^N$  et  $i \in \llbracket 1, N \rrbracket$ , si  $\|\tilde{y} - \tilde{x}\| < \epsilon$ , alors

$$\frac{1}{2} \max\{\sigma_i(d(\tilde{\phi}^n)_{\tilde{x}}), 1\} \leq \max\{\sigma_i(d(\tilde{\phi}^n)_{\tilde{y}}), 1\} \leq 2 \max\{\sigma_i(d(\tilde{\phi}^n)_{\tilde{x}}), 1\} \quad (51)$$

$$\|\tilde{\phi}^n(\tilde{y}) - \tilde{\phi}^n(\tilde{x}) - d(\tilde{\phi}^n)_{\tilde{x}}(\tilde{y} - \tilde{x})\| \leq \sup_{\tilde{z} \in \mathbb{R}^N} \|d^2(\tilde{\phi}^n)_{\tilde{z}}\| \epsilon^2 \leq \epsilon. \quad (52)$$

Soit  $E$  un ensemble  $\epsilon$ -séparé<sup>122</sup> maximal de  $\mathbb{T}^N$ . En particulier, les boules fermées  $B(x, \frac{\epsilon}{2})$  pour  $x \in E$  sont deux à deux disjointes (sinon  $E$  ne serait pas  $\epsilon$ -séparé) et les boules

122. c'est-à-dire dont deux points distincts sont à distance strictement supérieure à  $\epsilon$

fermées  $B(x, \epsilon)$  pour  $x \in E$  recouvrent  $\mathbb{T}^N$  (sinon  $E$  ne serait pas maximal). Considérons les cellules de Voronoï ouvertes dans  $\mathbb{T}^N$  définies par  $E$

$$\forall x \in E, \quad C_x = \{y \in \mathbb{T}^N : d(y, x) < \min_{x' \in E, x' \neq x} d(y, x')\}.$$

Par la finitude de la mesure et la non dénombrabilité des translations, quitte à translater un peu  $E$  dans  $\mathbb{T}^N$ , la mesure  $\mu(\bigcup_{x \in E} \partial C_x)$  est nulle, et donc  $\alpha = \{C_x : x \in E\}$  est une partition mesurable finie (à ensembles de mesure nulle près) de  $\mathbb{T}^N$ .

**Affirmation 1 :** Si  $\epsilon$  est assez petit (dépendant de  $n$ ), pour tout  $x \in E$ , l'image  $\phi^n(C_x)$  rencontre au plus  $c_N \prod_{i=1}^N \max\{\sigma_i(M_n(x)), 1\}$  atomes de  $\alpha$ .

**Démonstration.** Puisque la partie  $E$  est  $\epsilon$ -séparée, pour tout  $x' \in E$ , la cellule  $C_{x'}$  est contenue dans la boule  $B(x', \epsilon)$ . Donc par la formule (52), pour tout  $y \in C_x$ , nous avons  $\phi^n y \in \phi^n x + \mathcal{P}(\mathcal{V}_\epsilon(M_n(x))(B_{\mathbb{R}^N}(0, \epsilon)))$ . Pour tout  $x' \in E$ , comme le diamètre de la cellule  $C_{x'}$  est au plus  $2\epsilon$ , si  $\phi^n(C_x)$  rencontre  $C_{x'}$ , alors la boule  $B(x', \frac{\epsilon}{2})$  rencontre l'image  $\mathcal{P}(\mathcal{V}_{3\epsilon}(M_n(x))(B_{\mathbb{R}^N}(0, \epsilon)))$ . Donc le lemme 5.9 donne le résultat.  $\square$

**Affirmation 2 :** Nous avons

$$H_\mu(\phi^{-n}\alpha \mid \alpha) \leq N \ln 2 + \ln c_N + \sum_{i=1}^d \int_{\mathbb{T}^N} \ln^+(\sigma_i(M_n(x))) d\mu(x).$$

**Démonstration.** Par la définition de l'entropie relative des partitions mesurables, puisque si  $a_1, \dots, a_p \in [0, 1]$  vérifient  $\sum_{i=1}^p a_i = 1$  alors  $\sum_{i=1}^p \varphi(a_i) \leq \ln p$  comme vu dans la démonstration de la proposition 5.1 (1) et par l'affirmation 1 ci-dessus, nous avons

$$\begin{aligned} H_\mu(\phi^{-n}\alpha \mid \alpha) &= \sum_{x \in E} \mu(C_x) \left( \sum_{x' \in E} \varphi\left(\frac{\mu(\phi^{-n}C_{x'} \cap C_x)}{\mu(C_x)}\right) \right) \\ &\leq \sum_{x \in E} \mu(C_x) \ln \left( c_N \prod_{i=1}^N \max\{\sigma_i(M_n(x)), 1\} \right) \\ &= \sum_{x \in E} \mu(C_x) \left( \ln c_N + \sum_{i=1}^d \ln^+(\sigma_i(M_n(x))) \right) \end{aligned}$$

Or par la formule (51), nous avons

$$\mu(C_x) \ln^+(\sigma_i(M_n(x))) \leq \int_{y \in C_x} (\ln 2 + \ln^+(\sigma_i(M_n(y)))) d\mu(y).$$

L'affirmation 2 en découle.  $\square$

**Lemme 5.10.** Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret et  $\alpha$  une partition mesurable. Alors  $h_\mu(\phi, \alpha) \leq H_\mu(\phi^{-1}\alpha \mid \alpha)$ .

**Démonstration.** Par les assertions (3) et (4) (ii) de la proposition 5.1 et par récurrence,

nous avons

$$\begin{aligned}
H\left(\bigvee_{i=1}^{n-1} \phi^{-i}\alpha \mid \alpha\right) &= H\left(\bigvee_{i=2}^{n-1} \phi^{-i}\alpha \mid \alpha \vee \phi^{-1}\alpha\right) + H(\phi^{-1}\alpha \mid \alpha) \\
&\leq H\left(\bigvee_{i=2}^{n-1} \phi^{-i}\alpha \mid \phi^{-1}\alpha\right) + H(\phi^{-1}\alpha \mid \alpha) \\
&= H\left(\bigvee_{i=1}^{n-2} \phi^{-i}\alpha \mid \alpha\right) + H(\phi^{-1}\alpha \mid \alpha) \\
&\leq (n-1)H(\phi^{-1}\alpha \mid \alpha).
\end{aligned}$$

Donc  $H\left(\bigvee_{i=0}^{n-1} \phi^{-i}\alpha\right) = H\left(\bigvee_{i=1}^{n-1} \phi^{-i}\alpha \mid \alpha\right) + H(\alpha) \leq (n-1)H(\phi^{-1}\alpha \mid \alpha) + H(\alpha)$  et en divisant par  $n$  et en prenant la limite, nous avons le résultat.  $\square$

**Conclusion de la démonstration du théorème 5.8 :** Par le lemme 5.10 et l'affirmation 2, nous avons

$$\begin{aligned}
\frac{1}{n} h_\mu(\phi^n, \alpha) &\leq \frac{1}{n} H_\mu(\phi^{-n}\alpha \mid \alpha) \\
&\leq \frac{1}{n} \ln(2^N c_N) + \frac{1}{n} \sum_{i=1}^d \int_{\mathbb{T}^N} \ln^+(\sigma_i(M_n(x))) d\mu(x).
\end{aligned}$$

Notons  $D_n$  le membre de droite de cette inégalité, qui ne dépend plus de  $\epsilon$ . Soit  $(\epsilon_k)_{k \in \mathbb{N}}$  une suite strictement positive qui tend vers 0. En prenant des parties  $\epsilon_k$ -séparées maximales croissantes  $E_k$ , nous construisons comme ci-dessus une suite de partition  $(\alpha_k)_{k \in \mathbb{N}}$  de plus en plus fines dont le diamètre des atomes tend vers 0, telles que  $\frac{1}{n} h_\mu(\phi^n, \alpha_k) \leq D_n$  pour tout  $k \in \mathbb{N}$ . Par la proposition 5.7 (3) et le commentaire qui le suit, nous avons donc  $\frac{1}{n} h_\mu(\phi^n) = \lim_{k \rightarrow +\infty} \frac{1}{n} h_\mu(\phi^n, \alpha_k) \leq D_n$ . En faisant tendre  $n$  vers  $+\infty$ , le résultat en découle.  $\square$

## 5.4 Exercices

**Exercice E.27.** Soient  $(X, \mathcal{B}, \mu)$  un espace de probabilité et  $\alpha, \beta$  deux partitions mesurables de  $X$ . Montrer que les assertions suivantes sont équivalentes.

- (1)  $\alpha \perp \beta$  (les partitions mesurables  $\alpha$  et  $\beta$  sont indépendantes).
- (2)  $I_{\alpha \vee \beta} = I_\alpha + I_\beta$ .
- (3)  $H(\alpha \vee \beta) = H(\alpha) + H(\beta)$ .
- (4)  $I_{\alpha|\beta} = I_\alpha$ .
- (5)  $H(\alpha|\beta) = H(\alpha)$ .

**Exercice E.28.** Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret et  $\alpha$  une partition mesurable de  $(X, \mathcal{B})$ . Montrer que la suite  $(c_n = H_\mu(\alpha \mid \bigvee_{i=1}^n \phi^{-i}\alpha))_{n \in \mathbb{N}}$  est décroissante et converge vers  $h_\mu(\phi, \alpha)$ .

**Exercice E.29.** Pour tout  $\theta_0 \in \mathbb{R}/\mathbb{Z}$ , considérons la rotation  $\phi : \theta \mapsto \theta + \theta_0$  du cercle  $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$  d'angle  $\theta_0$ . Calculer l'entropie mesurée de  $\phi$  pour la mesure de Haar  $d\theta$  sur  $\mathbb{T}^1$ .

Montrer que si  $\phi$  est un homéomorphisme du cercle  $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$ , alors pour toute mesure de probabilité  $\phi$ -invariante sur  $\mathbb{T}^1$ , nous avons  $h_\mu(\phi) = 0$ .

**Exercice E.30.** Pour tout  $N$  dans  $\mathbb{N}$ , considérons la transformation  $\phi_N : \theta \mapsto N\theta$  du cercle  $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$ . Calculer l'entropie mesurée de  $\phi_N$  pour la mesure de Haar  $d\theta$  sur  $\mathbb{T}^1$ .

**Exercice E.31.** Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé inversible à temps discret et  $\alpha$  une partition mesurable de  $X$ . Montrer que  $h_\mu(\phi^{-1}, \alpha) = h_\mu(\phi, \alpha)$ . En déduire que

$$h_\mu(\phi^{-1}) = h_\mu(\phi) .$$

**Exercice E.32.** Soient  $(X, \mathcal{B}, \mu, \phi)$  et  $(Y, \mathcal{C}, \nu, \psi)$  deux systèmes dynamiques probabilisés à temps discret.

(1) Si  $(Y, \mathcal{C}, \nu, \psi)$  est un facteur de  $(X, \mathcal{B}, \mu, \phi)$ , montrer que

$$h_\nu(\psi) \leq h_\mu(\phi) .$$

(2) Si  $(X \times Y, \mathcal{B} \otimes \mathcal{C}, \mu \otimes \nu, \phi \times \psi)$  est le système dynamique probabilisé produit, montrer que

$$h_{\mu \otimes \nu}(\phi \times \psi) = h_\mu(\phi) + h_\nu(\psi) .$$

**Exercice E.33.** Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé inversible et  $\alpha$  une partition mesurable de  $X$ , telle que la suite de partitions  $(\alpha_n = \bigvee_{i=0}^n \phi^{-i} \alpha)_{n \in \mathbb{N}}$  soit asymptotiquement génératrice (c'est-à-dire telle que  $\bigcup_{n \in \mathbb{N}} \alpha_n$  engendre la  $\sigma$ -algèbre  $\mathcal{B}$ ). Montrer que

$$h_\mu(\phi) = 0 .$$

**Exercice E.34. (Théorème d'Abramov)** Soient  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret et  $A$  une partie mesurable de mesure non nulle de  $X$ . Notons  $(A, \mathcal{B}_A, \mu_A, \phi_A)$  le système dynamique induit (voir l'exercice E.13). Montrer que

$$\mu(A) h_{\mu_A}(\phi_A) = h_\mu(\phi) .$$

**Exercice E.35. (Facteur de Pinsker <sup>120</sup>)** Soit  $(X, \mathcal{B}, \mu, \phi)$  un système dynamique probabilisé à temps discret. Posons (avec la convention que  $\{B, {}^c B\}$  est la partition triviale  $\{X\}$  si  $B = \emptyset$  ou si  $B = X$ )

$$\Pi(\phi) = \{B \in \mathcal{B} : h_\mu(\phi, \{B, {}^c B\}) = 0\} .$$

Montrer que  $\Pi(\phi)$  est une  $\sigma$ -algèbre  $\phi$ -invariante. En déduire que  $(X, \Pi(\phi), \mu|_{\Pi(\phi)}, \phi)$  est un facteur d'entropie nulle de  $(X, \mathcal{B}, \mu, \phi)$  qui est maximal (au sens que toute semi-conjugaison mesurée de  $(X, \mathcal{B}, \mu, \phi)$  à valeurs dans un système dynamique probabilisé à temps discret d'entropie nulle est  $\Pi(\phi)$ -mesurable).

**Exercice E.36. (K-systèmes)** Un système dynamique probabilisé à temps discret  $(X, \mathcal{B}, \mu, \phi)$  est appelé un *K-système* (ou *système de Kolmogorov*) s'il existe une  $\sigma$ -algèbre  $\mathcal{C}$  contenue dans  $\mathcal{B}$  vérifiant

- $\phi^{-1}(\mathcal{C}) \subset \mathcal{C}$ ,
- $\bigcap_{n \geq 0} \phi^{-n}(\mathcal{C}) = \{\emptyset, X\}$  et
- $\bigvee_{n \in \mathbb{N}} \phi^n(\mathcal{C}) = \mathcal{B}$ .

Montrer que tout système de Bernoulli est un K-système. Montrer que pour tout K-système, toute partition mesurable  $\alpha$  non triviale modulo 0 et d'entropie finie vérifie  $h_\mu(\phi, \alpha) > 0$ . Un système dynamique probabilisé à temps discret vérifiant cette condition est dit *d'entropie complètement positive*.

## 5.5 Indications pour la résolution des exercices

**Exercice E.27.** Ceci a été démontré lors de la démonstration de l'assertion (4) (ii) de la proposition 5.1.

**Exercice E.28.** La décroissance de la suite  $(c_n)_{n \in \mathbb{N}}$  résulte de la proposition 5.1 (4) (ii). Notons  $c$  sa limite et posons  $a_n = H(\bigvee_{k=0}^{n-1} \phi^{-k} \alpha)$  pour tout  $n \geq 1$ . La proposition 5.1 (3) donne par récurrence

$$a_n = H\left(\bigvee_{k=0}^{n-1} \phi^{-k} \alpha\right) = \sum_{k=0}^{n-1} H(\phi^{-k} \alpha \mid \bigvee_{i=k+1}^{n-1} \phi^{-i} \alpha) + H(\phi^{-n+1} \alpha) = H(\alpha) + \sum_{k=0}^{n-1} c_{n-k-1}.$$

Par le lemme de Césaro, la suite  $a_n/n$  converge donc vers  $c$  et nous avons bien  $h(\phi, \alpha) = c$ .

**Exercice E.29.** Si  $\theta_0 = \frac{p}{q} \bmod \mathbb{Z}$  est rationnel, alors, pour toute partition mesurable  $\alpha$  de  $\mathbb{R}/\mathbb{Z}$ , nous avons

$$h_{d\theta}(\phi, \alpha) = \lim_{n \rightarrow \infty} \frac{1}{qn} H_{d\theta}\left(\bigvee_{i=0}^{nq-1} \phi^{-i} \beta\right) = \lim_{n \rightarrow \infty} \frac{1}{qn} H_{d\theta}\left(\bigvee_{i=0}^{q-1} \phi^{-i} \beta\right) = 0.$$

Donc  $h_{d\theta}(\phi) = 0$ .

Si  $\theta_0$  est irrationnel, alors il est élémentaire de montrer que la partition mesurable  $\alpha = \{[0, \frac{1}{2}[, [\frac{1}{2}, 1[ \}$  est génératrice et même que la suite de partitions  $\bigvee_{i=1}^n \phi^{-i} \alpha$  engendre la  $\sigma$ -algèbre des boréliens du cercle. Alors par le théorème de Kolmogorov-Sinai 5.3 (8), par l'exercice E.28 et par la proposition 5.1 (8), nous avons

$$h_{d\theta}(\phi) = h_{d\theta}(\phi, \alpha) = \lim_{n \rightarrow +\infty} H_{d\theta}\left(\alpha, \bigvee_{i=1}^n \phi^{-i} \alpha\right) = 0.$$

**Correction de l'exercice E.30.** Si  $N \geq 2$ , en considérant l'écriture N-adique, le système dynamique probabilisé  $(\mathbb{T}^1, d\theta, \phi_N)$  est conjugué (au sens de la mesure) au décalage sur  $\mathcal{A}^{\mathbb{N}}$  pour la mesure produit, avec  $\mathcal{A} = \{0, \dots, N-1\}$  muni de l'équiprobabilité (voir la correction de l'exercice E.23 (3)). Par l'invariance par conjugaison de l'entropie mesurée (voir la proposition 5.3 (5)) et l'exemple à la fin de la partie 5.2, nous avons

$$h_{d\theta}(\phi_N) = \log N.$$

**Correction de l'exercice E.31.** Pour tout  $n \in \mathbb{N}$ , nous avons  $H_\mu(\bigvee_{i=-n}^n \phi^{-i} \alpha) = H_\mu(\bigvee_{i=-n}^n (\phi^{-1})^{-i} \alpha)$ . En divisant par  $n$  et en prenant la limite (qui existe par le cas inversible de l'assertion (1) de la proposition 5.3), nous avons  $h_\mu(\phi, \alpha) = h_\mu(\phi^{-1}, \alpha)$ . Le résultat en découle en prenant la borne supérieure sur les partitions mesurables  $\alpha$ .

**Correction de l'exercice E.32.** (1) La démonstration est analogue à celle de la proposition 5.3 (5). Soit  $\theta : X \rightarrow X'$  une application (quitte à enlever des ensembles de mesure nulle) mesurable et préservant la mesure, telle que  $\theta \circ \phi = \phi' \circ \theta$ . Si  $\alpha'$  est une partition mesurable de  $X'$ , alors  $\theta^{-1} \alpha' = \{\theta^{-1}(A) : A \in \alpha'\}$  est une partition mesurable de  $X$ . La définition de l'entropie des partitions  $H_\mu$  et le fait que  $\theta_* \mu = \mu'$  impliquent que

$$H_{\mu'}\left(\bigvee_{i=0}^{n-1} (\phi')^{-i} \alpha'\right) = H_\mu\left(\bigvee_{i=0}^{n-1} \phi^{-i} \theta^{-1} \alpha'\right).$$

En divisant par  $n$  et en prenant la limite quand  $n$  tend  $+\infty$ , nous en déduisons que  $h_{\mu'}(\phi', \alpha') = h_{\mu}(\phi, \theta^{-1}\alpha')$ . En prenant la borne supérieure sur  $\alpha'$ , nous obtenons donc que  $h_{\mu'}(\phi') \leq h_{\mu}(\phi_{\mu})$ .

(2) Soient  $\alpha$  et  $\alpha'$  des partitions mesurables de  $X$ ,  $\beta$  et  $\beta'$  des partitions mesurables de  $Y$ , et  $k \in \mathbb{N}$ . L'ensemble  $\alpha \times \beta = \{A \times B : A \in \alpha, B \in \beta\}$  est une partition mesurable de  $X \times Y$ . Nous avons

$$(\phi \times \psi)^{-k}(\alpha \times \beta) = (\phi^{-k}\alpha) \times (\psi^{-k}\beta) \quad \text{et} \quad (\alpha \vee \alpha') \times (\beta \vee \beta') = (\alpha \times \beta) \vee (\alpha' \times \beta').$$

Il est élémentaire de vérifier que  $H_{\mu \otimes \nu}(\alpha \times \beta) = H_{\mu}(\alpha) + H_{\nu}(\beta)$ , donc

$$H_{\mu \otimes \nu}\left(\bigvee_{i=0}^{n-1} (\phi \times \psi)^{-k}(\alpha \times \beta)\right) = H_{\mu}\left(\bigvee_{i=0}^{n-1} \phi^{-i}\alpha\right) + H_{\nu}\left(\bigvee_{i=0}^{n-1} \psi^{-i}\beta\right).$$

En divisant par  $n$  et en passant à la limite, nous avons

$$h_{\mu \otimes \nu}(\phi \times \psi, \alpha \times \beta) = h_{\mu}(\phi, \alpha) + h_{\nu}(\psi, \beta).$$

Donc  $h_{\mu}(\phi, \alpha) + h_{\nu}(\psi, \beta) \leq h_{\mu \otimes \nu}(\phi \times \psi)$  et en prenant les bornes supérieures sur  $\alpha$  et sur  $\beta$ , nous avons  $h_{\mu}(\phi) + h_{\nu}(\psi) \leq h_{\mu \otimes \nu}(\phi \times \psi)$ . D'autre part, la formule centrée ci-dessus implique aussi que  $h_{\mu \otimes \nu}(\phi \times \psi, \alpha \times \beta) \leq h_{\mu}(\phi) + h_{\nu}(\psi)$ , et comme la  $\sigma$ -algèbre produit de  $X \times Y$  est engendrée par les parties mesurables qui sont des produits, nous avons (voir par exemple la proposition 5.7 (2) et (3))

$$h_{\mu \otimes \nu}(\phi \times \psi) = \sup_{\alpha, \beta} h_{\mu \otimes \nu}(\phi \times \psi, \alpha \times \beta) \leq h_{\mu}(\phi) + h_{\nu}(\psi),$$

où la borne supérieure est prise sur toutes les partitions mesurables  $\alpha$  de  $X$  et  $\beta$  de  $Y$  respectivement. D'où le résultat.

**Correction de l'exercice E.33.** Remarquer que la suite de partitions  $(\bigvee_{i=1}^n \phi^{-i}\alpha)_{n \in \mathbb{N}}$  engendre aussi  $\mathcal{B}$ , et étudier la suite  $H_{\mu}(\alpha | \bigvee_{i=1}^n \phi^{-i}\alpha)$ .

**Correction de l'exercice E.35.** Il est immédiat que l'ensemble  $\Pi(\phi)$  contient  $X$  (car si  $\alpha = \{X\}$ , alors  $\bigvee_{i=1}^n \phi^{-i}\alpha = \{X\}$  et  $H_{\mu}(\{X\}) = 0$ ) et est stable par passage au complémentaire.

Montrons que  $\Pi(\phi)$  est stable par union finie. Soient  $A, B \in \Pi(\phi)$  (que nous pouvons supposer différents de  $X, \emptyset$ ). Posons  $\alpha = \{A, {}^cA\}$  et  $\beta = \{B, {}^cB\}$ , et remarquons que  $\gamma = \{A \cup B, {}^c(A \cup B)\}$  est une partition moins fine que la partition jointe  $\alpha \vee \beta$ . Alors par la monotonie de l'entropie relative (voir la proposition 5.3 (3)), puis par la sous-additivité de l'entropie relative (voir la proposition 5.3 (2)), nous avons

$$0 \leq h_{\mu}(\phi, \gamma) \leq h_{\mu}(\phi, \alpha \vee \beta) \leq h_{\mu}(\phi, \alpha) + h_{\mu}(\phi, \beta) = 0.$$

Donc  $A \cup B \in \Pi(\phi)$ .

Montrons que  $\Pi(\phi)$  est stable par union croissante. Soit  $(A_n)_{n \in \mathbb{N}}$  une suite croissante (la partie  $A_n$  est contenue dans  $A_{n+1}$  pour tout  $n \in \mathbb{N}$ ) dans  $\Pi(\phi)$ . Nous pouvons supposer que ses éléments sont différents de  $X, \emptyset$ . Posons  $A = \bigcup_{n \in \mathbb{N}} A_n$ ,  $\alpha_n = \{A_n, {}^cA_n\}$  et  $\alpha = \{A, {}^cA\}$ . Par l'inégalité de Rokhlin (proposition 5.3 (4)), nous avons

$$0 \leq h_{\mu}(\phi, \alpha) = h_{\mu}(\phi, \alpha) - h_{\mu}(\phi, \alpha_n) \leq H_{\mu}(\alpha | \alpha_n).$$



Nous avons  $A_n \subset A$  et  $\lim_{n \rightarrow +\infty} \mu(A \Delta A_n) = \lim_{n \rightarrow +\infty} \mu(A \setminus A_n) = 0$  car la mesure est finie et de même  $\lim_{n \rightarrow +\infty} \mu({}^c A \Delta {}^c A_n) = \lim_{n \rightarrow +\infty} \mu({}^c A_n \setminus {}^c A) = 0$  car  ${}^c A$  est l'intersection décroissante  $\bigcap_{n \in \mathbb{N}} {}^c A_n$ . Donc par la continuité de l'entropie des partitions (voir la proposition 5.1 (7)), nous avons  $\lim_{n \rightarrow +\infty} H_\mu(\alpha | \alpha_n) = 0$ . Par conséquent  $h_\mu(\phi, \alpha) = 0$  et  $A \in \Pi(\phi)$ .

Montrons que  $\Pi(\phi)$  est  $\phi$ -invariante. Soit  $A \in \Pi(\phi)$  (que nous pouvons supposer différente de  $X, \emptyset$ ). Posons  $\alpha = \{A, {}^c A\}$ . Comme la partition mesurable  $\bigvee_{k=1}^n \phi^{-k} \alpha$  est moins fine que la partition mesurable  $\bigvee_{k=0}^n \phi^{-k} \alpha$  et par la monotonie de l'entropie des partitions (voir la proposition 5.1 (4) (i)), nous avons

$$\frac{1}{n} H_\mu\left(\bigvee_{i=1}^n \phi^{-i} \alpha\right) \leq \frac{n+1}{n} \frac{1}{n+1} H_\mu\left(\bigvee_{i=0}^n \phi^{-i} \alpha\right).$$

Le membre de gauche de cette inégalité converge vers  $h_\mu(\phi, \phi^{-1} \alpha) \geq 0$  et celui de droite vers  $h_\mu(\phi, \alpha) = 0$ . Donc  $h_\mu(\phi, \phi^{-1} \alpha) = 0$  et  $\phi^{-1}(A) \in \Pi(\phi)$ .

L'application  $\text{id} : X \rightarrow X$  est mesurable de  $(X, \mathcal{B})$  dans  $(X, \Pi(\phi))$ , car  $\Pi(\phi)$  est contenue dans  $\mathcal{B}$ . Elle envoie la mesure  $\mu : \mathcal{B} \rightarrow [0, 1]$  sur la restriction  $\mu|_{\Pi(\phi)}$  de  $\mu$  à  $\Pi(\phi)$ . C'est donc une semi-conjugaison mesurée du système dynamique probabilisé  $(X, \mathcal{B}, \mu, \phi)$  à valeurs dans le système dynamique probabilisé  $(X, \Pi(\phi), \mu|_{\Pi(\phi)}, \phi)$ , et celui-ci est un facteur de  $(X, \mathcal{B}, \mu, \phi)$ .

Montrons que l'entropie de  $(X, \Pi(\phi), \mu|_{\Pi(\phi)}, \phi)$  est nulle. Pour toute partition mesurable  $\alpha = \{A_1, \dots, A_n\}$  pour la  $\sigma$ -algèbre  $\Pi(\phi)$ , si  $\alpha_i = \{A_i, {}^c A_i\}$  alors  $\alpha = \bigvee_{i=1}^n \alpha_i$ . Donc par la sous-additivité de l'entropie relative (voir la proposition 5.3 (2)), nous avons

$$0 \leq h_\mu(\phi, \alpha) \leq \sum_{i=1}^n h_\mu(\phi, \alpha_i) = 0.$$

D'où l'annulation de l'entropie de  $(X, \Pi(\phi), \mu|_{\Pi(\phi)}, \phi)$  en passant à la borne supérieure sur les partitions  $\Pi(\phi)$ -mesurables  $\alpha$ .

Si  $(X', \mathcal{B}', \mu', \phi')$  est un facteur de  $(X, \mathcal{B}, \mu, \phi)$  d'entropie nulle, soient  $\theta : X \rightarrow X'$  une semi-conjugaison (quitte à enlever des parties mesurables invariantes de mesures nulles), et  $B = \theta^{-1}(B')$  un élément de la  $\sigma$ -algèbre  $\theta^{-1}(\mathcal{B}')$ . Comme dans la démonstration de l'exercice E.32 (1), posons  $\alpha' = \{B', {}^c B'\}$ , qui est une partition mesurable de  $X'$ . Alors  $\theta^{-1} \alpha' = \{B, {}^c B\}$  est une partition mesurable de  $X$ . La définition de l'entropie des partitions  $H_\mu$  et le fait que  $\theta_* \mu = \mu'$  impliquent que

$$H_\mu\left(\bigvee_{i=0}^{n-1} \phi^{-i} \theta^{-1} \alpha'\right) = H_{\mu'}\left(\bigvee_{i=0}^{n-1} (\phi')^{-i} \alpha'\right).$$

En divisant par  $n$  et en prenant la limite quand  $n$  tend  $+\infty$ , nous en déduisons que

$$0 \leq h_\mu(\phi, \{B, {}^c B\}) = h_\mu(\phi, \theta^{-1} \alpha') = h_{\mu'}(\phi', \alpha') \leq h_{\mu'}(\phi') = 0.$$

Donc  $B \in \Pi(\phi)$ , ce qu'il fallait démontrer.

## 6 Dynamique symbolique markovienne

Nous étudions dans cette partie les propriétés de dynamique mesurée des systèmes dynamiques symboliques, c'est-à-dire des sous-systèmes dynamiques des systèmes de Bernoulli  $(\mathcal{A}^{\mathbb{N}}, \sigma_+)$  ou  $(\mathcal{A}^{\mathbb{Z}}, \sigma)$  sur un alphabet  $\mathcal{A}$  (définis dans l'exemple (2) de la partie 0.1). Nous renvoyons à la partie 11.3.1 pour une étude des propriétés de dynamique topologique. Un point clef (voir le lemme 6.3) est l'abondance des mesures invariantes qu'ils possèdent. Nous renvoyons à la partie 11 pour l'utilisation de systèmes dynamiques symboliques pour la description et la compréhension de systèmes dynamiques topologiques ou différentiables.

### 6.1 Chaînes de Markov<sup>126</sup> sur un alphabet fini

Le but de cette sous-partie est de donner une construction d'une famille de mesures invariantes sur des systèmes dynamiques symboliques, classique en théorie des probabilités, à l'aide de matrices de transition (ou stochastiques), de calculer leur entropie mesurée et d'étudier leur propriété de mélange.

#### Matrices à coefficients positifs ou nuls

Afin d'étudier le problème d'unicité des mesures stationnaires que nous définirons ci-dessous, nous aurons besoin de résultats classiques sur les matrices à coefficients positifs ou nuls. Ces résultats serviront aussi dans la partie 11.3.1.

Une matrice carrée  $B = (b_{ij})_{i,j \in I}$  à coefficients positifs ou nuls est dite

- *irréductible* si pour tous les  $i, j$  dans  $I$ , il existe  $n > 0$  tel que  $(B^n)_{ij} \neq 0$ ,
- *apériodique*<sup>123</sup> s'il existe un entier  $k \in \mathbb{N} - \{0\}$  tel que les coefficients de  $B^k$  soient tous strictement positifs.<sup>124</sup>

La condition d'apériodicité est plus forte que celle d'irréductibilité.<sup>125</sup> Voir les exercices E.66 et E.38.

Nous aurons besoin du résultat classique suivant.

**Théorème 6.1. (Théorème de Perron<sup>126</sup>-Frobenius<sup>126</sup>)** *Soit  $B$  une matrice carrée à coefficients réels positifs ou nuls, qui est apériodique.*

- a) *La matrice  $B$  admet un vecteur propre  $v_{\max}$  à coefficients strictement positifs. Tout vecteur propre à coefficients positifs ou nuls est proportionnel à  $v_{\max}$ .*

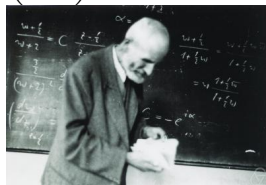
123. Voir l'exercice E.38 pour une explication de la terminologie. Certains ouvrages utilisent la terminologie « transitive », gardant la terminologie « apériodique » pour être de période 1 au sens de l'exercice E.38.

124. Noter la différence formelle dans l'ordre des quantificateurs avec le fait d'être irréductible : l'apériodicité signifie qu'il existe  $n > 0$  tel que pour tous les  $i, j$  dans  $I$ , nous avons  $(B^n)_{ij} \neq 0$ .

125. Par exemple,  $B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  est irréductible, mais pas apériodique.



126. Andrei Markov  
1856-1922



Oskar Perron  
1880-1975



Ferdinand Frobenius  
1849-1917



Carl Gauss  
1777-1855

b) Si  $\lambda_{\max} = \lambda_{\max}(B)$  est la valeur propre de  $B$  correspondant à  $v_{\max}$ , alors  $\lambda_{\max}$  est simple, strictement positive, et toute autre valeur propre complexe  $\lambda$  de  $B$  vérifie

$$|\lambda| < \lambda_{\max} .$$

**Démonstration.** Soit  $k \in \mathbb{N} \setminus \{0\}$  tel que les coefficients de  $B^k$  soient tous strictement positifs. Montrons tout d'abord que quitte à remplacer  $B$  par  $B^k$ , nous pouvons supposer que  $k = 1$ .

En effet, soit  $v_{\max}$  un vecteur propre de  $B^k$  à coefficients strictement positifs tel que tout vecteur propre de  $B^k$  à coefficients positifs ou nuls lui soit proportionnel, et dont la valeur propre  $\lambda_{\max}$  soit simple et strictement plus grande que les valeurs absolues des autres valeurs propres de  $B^k$ . Alors  $\sqrt[k]{\lambda_{\max}}$  est une valeur propre simple de  $B$ , strictement plus grande que les valeurs absolues des autres valeurs propres de  $B$ . De plus  $Bv_{\max}$ , qui est un vecteur propre de  $B^k$  pour la valeur propre  $\lambda_{\max}$ , est proportionnel à  $v_{\max}$ . Donc  $v_{\max}$  est un vecteur propre de  $B$ , et sa valeur propre est nécessairement  $\sqrt[k]{\lambda_{\max}}$ . Si  $v$  est un vecteur propre de  $B$  à coefficients positifs ou nuls, alors il l'est aussi pour  $B^k$ , donc est proportionnel à  $v_{\max}$ .

a) Soit  $q \in \mathbb{N} \setminus \{0\}$  tel que  $B \in \mathcal{M}_q(\mathbb{R})$ . Considérons le convexe compact

$$C = \{v = (x_1, \dots, x_q) \in \mathbb{R}^q : x_i \geq 0 \text{ et } x_1 + x_2 + \dots + x_q = 1\} .$$

Puisque la matrice  $B$  est à coefficients positifs ou nuls, l'application linéaire  $B : \mathbb{R}^q \rightarrow \mathbb{R}^q$  associée induit une application  $f_B : C \rightarrow C$  uniquement définie par  $\mathbb{R}f_B(v) = \mathbb{R}Bv$  pour tout  $v$  dans  $C$ .

Munissons  $C$  de la distance (dite *distance de Hilbert*)

$$d(v, w) = \ln [a, b, v, w]$$

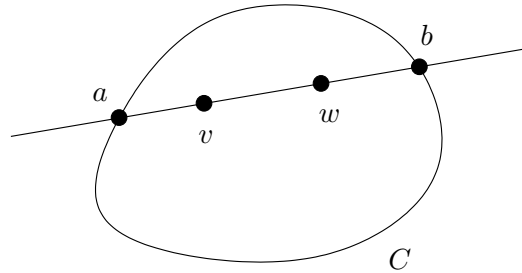
où  $a, b$  sont les points du bord de  $C$  sur la droite  $\mathcal{D}$  passant par  $v$  et  $w$  (supposés distincts), avec  $a$  du côté de  $v$ , et où  $[a, b, v, w]$  désigne le birapport de ces quatre points (rappelons que

$$[a, b, v, w] = \frac{(w - a)(b - v)}{(v - a)(b - w)} ,$$

en utilisant une (et n'importe laquelle) coordonnée affine sur la droite  $\mathcal{D}$ .)

(Le fait que  $d$  vérifie l'inégalité triangulaire peut être démontré aisément. Nous omettons cette vérification car ce fait n'est pas utilisé dans l'argument ci-dessous. Voir par exemple [Har].)

Puisque les coefficients de  $B$  sont tous strictement positifs, l'image  $f_B(C)$  est contenue dans l'intérieur relatif du convexe  $C$ .



L'application  $t \mapsto \frac{t+c}{t} = 1 + \frac{c}{t}$  est strictement décroissante sur  $]0, +\infty[$  pour toute constante  $c > 0$ . Donc en notant  $a'$  et  $b'$  les points d'intersection avec le bord de  $C$  de la droite passant par  $f_B(v)$  et  $f_B(w)$  orientée de  $f_B(v)$  vers  $f_B(w)$ , avec  $b'$  du côté de  $f_B(w)$ , et en posant  $c = f_B(w) - f_B(v)$ , nous avons  $\frac{b' - f_B(v)}{b' - f_B(w)} = \frac{b' - f_B(w) + c}{b' - f_B(w)} > \frac{f_B(b) - f_B(w) + c}{f_B(b) - f_B(w)}$ . Comme  $f_B$  conserve le birapport, nous avons donc, pour tous les  $v, w$  distincts dans  $C$ ,

$$d(f_B(v), f_B(w)) < d(v, w) .$$

Ceci montre que  $f_B$  contient un unique point fixe dans  $C$ , qui est dans l'intérieur relatif de  $C$  : c'est l'unique point  $v_0$  qui minimise la fonction continue  $v \mapsto d(v, f_B(v))$  sur le compact  $f_B(C)$ .

b) Supposons par l'absurde qu'il existe une valeur propre complexe  $\lambda$ , qui soit différente de  $\lambda_{\max}$  ou qui soit égale à  $\lambda$  mais de multiplicité au moins 2, telle que  $|\lambda| \geq \lambda_{\max}$ . En utilisant le théorème de Jordan et en discutant suivant le fait que  $\lambda$  est réelle ou non, nous pourrions trouver  $|\lambda| \geq \lambda_{\max}$  et un sous-espace vectoriel  $B$ -invariant  $V$  de dimension 2 ou 3 de  $\mathbb{R}^q$ , contenant  $v_{\max}$  et tel que dans une base de  $V$  ayant  $v_{\max}$  comme premier vecteur, la restriction de  $B$  soit donnée par une matrice de la forme

$$\begin{pmatrix} \lambda_{\max} & 0 \\ 0 & \pm|\lambda| \end{pmatrix}, \quad \begin{pmatrix} \lambda_{\max} & 1 \\ 0 & \pm|\lambda| \end{pmatrix} \quad \text{ou} \quad \begin{pmatrix} \lambda_{\max} & 0 & 0 \\ 0 & |\lambda| \cos \theta & -|\lambda| \sin \theta \\ 0 & |\lambda| \sin \theta & |\lambda| \cos \theta \end{pmatrix} .$$

Dans ces trois cas, il n'existe pas dans  $V$  de cône  $C'$  convexe fermé *saillant* (i.e. ne contenant pas de droite vectorielle), de sommet l'origine de  $V$ , tel que  $v_{\max}$  appartienne à l'intérieur de  $C'$  et dont l'image par  $B$  est contenue dans l'intérieur de  $C'$  (auquel nous avons rajouté 0). Ceci contredit ce que nous avons vu dans la démonstration de l'assertion a).  $\square$

### Matrices de transition

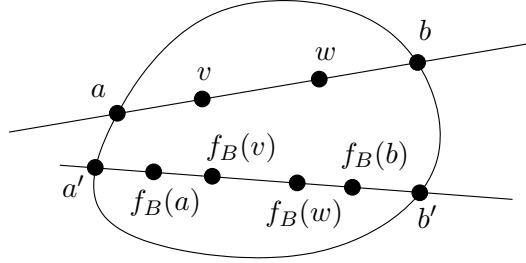
Soit  $\mathcal{A}$  un ensemble fini<sup>127</sup>, de cardinal au moins 2, muni de la topologie discrète, appelé un *alphabet*. Une *matrice de transition*<sup>128</sup> sur  $\mathcal{A}$  est par définition une matrice  $\pi = (\pi_{\alpha\beta})_{\alpha, \beta \in \mathcal{A}}$  à coefficients positifs ou nuls tels que la somme des coefficients de chaque ligne vaut 1 : pour tout  $\alpha$  dans  $\mathcal{A}$ , nous avons

$$\sum_{\beta \in \mathcal{A}} \pi_{\alpha\beta} = 1 .$$

La *matrice d'incidence associée* à la matrice de transition  $\pi$  est par définition la matrice  $A_\pi = (A_{\alpha\beta})_{\alpha, \beta \in \mathcal{A}}$  définie par  $A_{\alpha\beta} = \begin{cases} 0 & \text{si } \pi_{\alpha\beta} = 0 \\ 1 & \text{sinon} . \end{cases}$  Notons que  $\pi$  est apériodique (respectivement irréductible) si et seulement si  $A_\pi$  l'est.

127. Certains des résultats ci-dessous s'étendent au cas des alphabets dénombrables (discrets), en rajoutant parfois des hypothèses assez fortes sur les matrices de transition, voir par exemple [Kit], ainsi que [Sar1, Sar3, Sar4, BuzS, Sar5, BrPP] pour des développements.

128. Certains ouvrages disent aussi *matrice stochastique*, et dans [BenP] c'est la matrice transposée qui est appelée une matrice stochastique.



Une mesure (borélienne) de probabilité  $\nu$  sur l'espace discret  $\mathcal{A}$  est identifiée de manière usuelle avec le vecteur  $(\nu_\alpha)_{\alpha \in \mathcal{A}}$  de  $\mathbb{R}^{\mathcal{A}}$  défini par  $\nu_\alpha = \nu(\{\alpha\})$ , de sorte que pour toute partie  $E$  de  $\mathcal{A}$ , nous ayons  $\nu(E) = \sum_{\alpha \in E} \nu_\alpha$ . Notons que le vecteur  $(\nu_\alpha)_{\alpha \in \mathcal{A}}$  est à coefficients positifs ou nuls, de somme égale à 1

$$\sum_{\alpha \in \mathcal{A}} \nu_\alpha = 1,$$

et que nous l'identifions à la **matrice ligne** correspondante. Une *mesure  $\pi$ -stationnaire* (ou *mesure stationnaire* lorsque  $\pi$  est sous-entendue) sur  $\mathcal{A}$  est par définition une mesure de probabilité  $\nu$  sur  $\mathcal{A}$  qui est un vecteur propre de  ${}^t\pi$  pour la valeur propre 1. Avec  $\nu$  identifié à une matrice ligne comme dit ci-dessus, ceci revient à demander que l'égalité matricielle  $\nu\pi = \nu$  soit vérifiée, c'est-à-dire que, pour tout  $\beta$  dans  $\mathcal{A}$ , nous ayons

$$\sum_{\alpha \in \mathcal{A}} \nu_\alpha \pi_{\alpha\beta} = \nu_\beta.$$

**Lemme 6.2.** *Soit  $\mathcal{A}$  comme ci-dessus.*

- (1) *Toute matrice de transition  $\pi$  sur  $\mathcal{A}$  admet au moins une mesure  $\pi$ -stationnaire.*
- (2) *Si  $\pi$  est apériodique, alors une mesure  $\pi$ -stationnaire est unique, c'est l'unique vecteur  $\nu^\pi = (\nu_\alpha^\pi)_{\alpha \in \mathcal{A}}$  tel que  $\nu^\pi \pi = \nu^\pi$  et  $\sum_{\alpha \in \mathcal{A}} \nu_\alpha^\pi = 1$ . La mesure  $\nu^\pi$  est de support égal à  $\mathcal{A}$  (tous les coefficients de  $\nu^\pi$  sont strictement positifs). Le réel 1 est une valeur propre simple de  $\pi$  et les valeurs propres  $\lambda$  différentes de 1 vérifient  $|\lambda| < 1$ .*

**Démonstration.** (1) L'ensemble des matrices de transition est un fermé borné de  $\mathbb{R}^{\mathcal{A}^2}$ , donc compact. Pour tout  $n \in \mathbb{N}$ , il est immédiat de vérifier que les matrices  $\pi^n$  et  $\frac{1}{n+1} \sum_{k=0}^n \pi^k$  sont des matrices de transition. Soit  $Q$  un point d'accumulation de la suite  $(\frac{1}{n+1} \sum_{k=0}^n \pi^k)_{n \in \mathbb{N}}$ . Alors il est immédiat de vérifier que  $Q\pi = Q$ , donc chaque ligne de  $Q$  est une mesure  $\pi$ -stationnaire sur  $\mathcal{A}$ .

(2) La matrice  ${}^t\pi$  est aussi une matrice apériodique. Puisque  $\pi^n$  est aussi une matrice de transition pour tout  $n \in \mathbb{N}$ , la norme matricielle de  $({}^t\pi)^n = {}^t(\pi^n)$  est bornée. Donc le rayon spectral  $\rho({}^t\pi)$  de  ${}^t\pi$  est inférieur ou égal à 1, et en fait égal à 1 car 1 est une valeur propre de  ${}^t\pi$ . Le résultat découle alors du théorème de Perron-Frobenius 6.1 appliqué à la matrice  ${}^t\pi$ , puisque  $\lambda_{\max}({}^t\pi) = \rho({}^t\pi) = 1$ .  $\square$

### Mesures de Markov

Par une construction classique en théorie des probabilités, la donnée d'une matrice de transition  $\pi$  et d'une mesure  $\pi$ -stationnaire  $\nu$  sur  $\mathcal{A}$  définit une marche aléatoire  $(\omega_n)_{n \in \mathbb{N}}$  sur l'espace des états  $\mathcal{A}$ , de probabilité de passage de l'état  $\alpha \in \mathcal{A}$  à l'état  $\beta \in \mathcal{A}$  égale à  $\pi_{\alpha\beta}$ , et de loi initiale  $\nu$ , et donc une mesure de probabilité  $\mathbb{P}$  sur l'espace mesurable produit  $\mathcal{A}^{\mathbb{N}}$  des trajectoires de la marche aléatoire :

$$\forall \alpha, \beta \in \mathcal{A}, \quad \forall k \in \mathbb{N} \setminus \{0\}, \quad \mathbb{P}[\omega_0 = \alpha] = \nu_\alpha \quad \text{et} \quad \mathbb{P}[\omega_k = \beta \mid \omega_0 = \alpha] = (\pi^k)_{\alpha,\beta}.$$

Mais nous allons démontrer cela de manière autonome. Nous utilisons la convention usuelle qu'un produit vide de nombres positifs ou nuls est égal à 1.

**Lemme 6.3.** *Pour toute matrice de transition  $\pi$  et toute mesure  $\pi$ -stationnaire  $\nu$  sur  $\mathcal{A}$ , il existe une et une seule mesure de probabilité  $\mathbb{P}_{\pi,\nu}$  (respectivement  $\mathbb{P}_{\pi,\nu,+}$ ) sur l'espace mesurable produit  $\Sigma = \mathcal{A}^{\mathbb{Z}}$  (respectivement  $\Sigma_+ = \mathcal{A}^{\mathbb{N}}$ ) invariante par le décalage à gauche  $\sigma$  (respectivement  $\sigma_+$ ) telle que, pour tous les  $n \in \mathbb{N}$  et  $\alpha_0, \dots, \alpha_n \in \mathcal{A}$ ,*

$$\mathbb{P}_{\pi,\nu}(\{\omega \in \Sigma : \omega_0 = \alpha_0, \dots, \omega_n = \alpha_n\}) = \nu_{\alpha_0} \prod_{i=1}^n \pi_{\alpha_{i-1}, \alpha_i}.$$

(respectivement

$$\mathbb{P}_{\pi,\nu,+}(\{\omega \in \Sigma_+ : \omega_0 = \alpha_0, \dots, \omega_n = \alpha_n\}) = \nu_{\alpha_0} \prod_{i=1}^n \pi_{\alpha_{i-1}, \alpha_i}.$$

De plus, si <sup>129</sup> le support de la mesure  $\nu$  est égal à  $\mathcal{A}$ , alors le support de la mesure  $\mathbb{P}_{\pi,\nu}$  (respectivement  $\mathbb{P}_{\pi,\nu,+}$ ) est égal à  $\Sigma_{\pi} = \{(\alpha_i)_{i \in \mathbb{Z}} \in \mathcal{A}^{\mathbb{Z}} : \pi_{\alpha_i, \alpha_{i+1}} \neq 0\}$  (respectivement  $\Sigma_{\pi,+} = \{(\alpha_i)_{i \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}} : \pi_{\alpha_i, \alpha_{i+1}} \neq 0\}$ ).

Nous appellerons  $\mathbb{P}_{\pi,\nu}$  (respectivement  $\mathbb{P}_{\pi,\nu,+}$ ) la *mesure de Markov bilatère* sur  $\Sigma_{\pi}$  (respectivement *unilatère* sur  $\Sigma_{\pi,+}$ ) associée à  $(\pi, \nu)$ . L'application  $\sigma_{\pi} = \sigma|_{\Sigma_{\pi}}$  (respectivement  $\sigma_{\pi,+} = (\sigma_+)|_{\Sigma_{\pi,+}}$ ) est parfois appelée le *décalage de Markov* associé à  $(\pi, \nu)$ . Nous reviendrons dans la partie 11.3 à l'étude des systèmes dynamiques mesurés  $(\Sigma_{\pi}, \mathbb{P}_{\pi,\nu}, \sigma_{\pi})$  et  $(\Sigma_{\pi,+}, \mathbb{P}_{\pi,\nu,+}, \sigma_{\pi,+})$ , appelés les *systèmes de Markov bilatère* et *unilatères* respectivement associés à  $(\pi, \nu)$ .

**Exemple.** Les mesures de Bernoulli sont des cas particuliers de mesures de Markov. En effet, soit  $\nu$  une mesure de probabilité sur  $\mathcal{A}$ . Notons  $\pi = (\pi_{\alpha,\beta})_{\alpha,\beta \in \mathcal{A}}$  la matrice dont toutes les lignes sont égales à  $\nu$ , c'est-à-dire vérifiant  $\pi_{\alpha,\beta} = \nu_{\beta}$  pour tous les  $\alpha, \beta \in \mathcal{A}$ . Alors  $\nu$  est une mesure  $\pi$ -stationnaire car  $\sum_{\alpha \in \mathcal{A}} \nu_{\alpha} \pi_{\alpha,\beta} = \sum_{\alpha \in \mathcal{A}} \nu_{\alpha} \nu_{\beta} = \nu_{\beta}$  pour tout  $\beta \in \mathcal{A}$ . La mesure de Markov bilatère  $\mathbb{P}_{\pi,\nu}$  associée à  $(\pi, \nu)$  vérifie donc

$$\mathbb{P}_{\pi,\nu}(\{\omega \in \Sigma : \omega_0 = \alpha_0, \dots, \omega_n = \alpha_n\}) = \prod_{i=0}^n \nu_{\alpha_i}.$$

Par unicité, nous avons donc que  $\mathbb{P}_{\pi,\nu}$  est la mesure de Bernoulli bilatère  $\nu^{\mathbb{Z}}$ .

De même, la mesure de Markov unilatère  $\mathbb{P}_{\pi,\nu,+}$  est égale à la mesure de Bernoulli unilatère  $\nu^{\mathbb{N}}$ .

**Démonstration.** Nous ne considérons que le cas bilatère, le cas unilatère se traite de même.

En généralisant la terminologie introduite dans l'exemple (2) de la partie 0.1, nous appellerons *cylindre* dans  $\Sigma = \mathcal{A}^{\mathbb{Z}}$  toute partie de  $\Sigma$  de la forme

$$[\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_k = \{\omega \in \Sigma : \forall i \in \llbracket m, n \rrbracket, \omega_{i-m+k} = \alpha_i\},$$

avec  $k$  et  $m \leq n$  dans  $\mathbb{Z}$  (nous dirons que ce cylindre *part du rang  $k$*  et que sa *longueur* est  $n - m + 1$ ) et  $\alpha_m, \alpha_{m+1}, \dots, \alpha_n$  dans  $\mathcal{A}$ . Notons que pour toute mesure  $\mu$  sur  $\Sigma$  invariante par le décalage  $\sigma$ , et pour tout  $k \in \mathbb{Z}$ , nous avons

$$\mu([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_k) = \mu(\sigma^{-k}([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_0)) = \mu([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_0).$$

129. Voir l'exercice E.40 pour des conditions équivalentes à cette hypothèse

Notons que puisque  $\mathcal{A}$  est fini, l'ensemble  $\mathcal{E}_{\text{cyl}}$  des unions finies de cylindres est une algèbre de Boole engendrant la  $\sigma$ -algèbre produit de  $\Sigma$  (voir la note de bas de page 29 pour des détails).

Afin de construire une mesure de probabilité borélienne  $\mathbb{P}_{\pi,\nu}$  sur  $\Sigma$ , en procédant comme dans l'exemple (2) de la partie 0.2, nous la définissons sur tous les cylindres en posant (avec la convention usuelle qu'un produit vide est égal à 1), pour tous les  $k, m, n \in \mathbb{Z}$  tels que  $m \leq n$  et  $\alpha_m, \dots, \alpha_n$  dans  $\mathcal{A}$ ,

$$\mathbb{P}_{\pi,\nu}([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_k) = \nu_{\alpha_m} \prod_{i=m}^{n-1} \pi_{\alpha_i \alpha_{i+1}} \in [0, +\infty[. \quad (53)$$

Nous étendons par additivité cette application  $\mathbb{P}_{\pi,\nu}$  sur l'ensemble  $\mathcal{E}_{\text{cyl}}$  des unions finies disjointes de cylindres, ce qui ne pose pas de conflit lorsqu'un cylindre est une union disjointe de cylindres : en effet, par récurrence, puisque  $\nu\pi = \nu$  et puisque la somme des coefficients de toute ligne de la matrice  $\pi$  est égale à 1, nous avons

$$\begin{aligned} \mathbb{P}_{\pi,\nu}([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_m) &= \sum_{\alpha \in \mathcal{A}} \mathbb{P}_{\pi,\nu}([\alpha, \alpha_m, \dots, \alpha_n]_{m-1}) \\ &= \sum_{\alpha \in \mathcal{A}} \mathbb{P}_{\pi,\nu}([\alpha_m, \dots, \alpha_n, \alpha]_m). \end{aligned}$$

Il est alors immédiat de montrer que l'application  $\mathbb{P}_{\pi,\nu} : \mathcal{E}_{\text{cyl}} \rightarrow [0, +\infty[$  vérifie les propriétés i), ii) et iii) du théorème de Carathéodory 0.1. En effet, la propriété ii) est acquise par construction. Pour montrer la propriété i), ainsi que le fait que l'extension soit une mesure de probabilité, il suffit de remarquer que puisque  $\nu$  est une mesure de probabilité sur  $\mathcal{A}$ , nous avons

$$\mathbb{P}_{\pi,\nu}(X) = \mathbb{P}_{\pi,\nu}\left(\bigsqcup_{\alpha \in \mathcal{A}} [\alpha]_0\right) = \sum_{\alpha \in \mathcal{A}} \mathbb{P}_{\pi,\nu}([\alpha]_0) = \sum_{\alpha \in \mathcal{A}} \nu_{\alpha} = 1.$$

Enfin, la condition iii) du théorème de Carathéodory est automatiquement satisfaite, car les unions finies de cylindres sont des compacts de  $\Sigma$ , donc si  $(C_n)_{n \in \mathbb{N}}$  est une suite décroissante dans  $\mathcal{E}_{\text{cyl}}$  telle que  $\bigcap_{n \in \mathbb{N}} C_n = \emptyset$ , alors il existe  $n_0$  tel que, pour tout  $n \geq n_0$ , nous avons  $C_n = \emptyset$ .

La condition d'invariance de  $\mathbb{P}_{\pi,\nu}$  par le décalage est satisfaite par l'invariance des mesures des cylindres par décalage (le terme de droite de l'égalité (53) est indépendant de  $k$ ) et par l'unicité dans le théorème de Carathéodory (voir aussi la Proposition 0.2 (3)).

Explicitons pour conclure le support de  $\mathbb{P}_{\pi,\nu}$ . Pour tous les  $\alpha_m, \alpha_{m+1}, \dots, \alpha_n \in \mathcal{A}$ , nous avons  $\mathbb{P}_{\pi,\nu}([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_m) \neq 0$  si et seulement si  $\pi_{\alpha_i \alpha_{i+1}} \neq 0$  pour tout  $i \in \llbracket m, n-1 \rrbracket$ , car  $\nu_{\alpha} \neq 0$  pour tout  $\alpha \in \mathcal{A}$ , vu que le support de  $\nu$  est  $\mathcal{A}$ . Puisque les cylindres forment une base d'ouverts de la topologie produit de  $X$ , le résultat en découle.  $\square$

Le résultat suivant calcule l'entropie mesurée des systèmes de Markov bilatère  $(\Sigma, \mathbb{P}_{\pi,\nu}, \sigma)$  et unilatère  $(\Sigma_+, \mathbb{P}_{\pi,\nu,+}, \sigma_+)$ .

**Proposition 6.4.** *Soient  $\pi$  une matrice de transition et  $\nu$  une mesure  $\pi$ -stationnaire sur  $\mathcal{A}$ . Alors*

$$h_{\mathbb{P}_{\pi,\nu}}(\sigma) = h_{\mathbb{P}_{\pi,\nu,+}}(\sigma_+) = - \sum_{\alpha, \beta \in \mathcal{A}} \nu_{\alpha} \pi_{\alpha\beta} \ln(\pi_{\alpha\beta}).$$

**Démonstration.** Nous ne considérons que le cas bilatère, le cas unilatère se traite de même.

Reprenons les notations de la démonstration précédente concernant les cylindres. Considérons la partition mesurable  $\xi_0 = \{[\alpha]_0 : \alpha \in \mathcal{A}\}$  de  $\mathcal{A}^{\mathbb{Z}}$  par les cylindres de longueur 1 partant en 0. Comme nous l'avons vu dans le calcul (49) de l'entropie des systèmes de Bernoulli avec leur mesure de Bernoulli dans la partie 5.2, cette partition est génératrice pour la transformation inversible  $\sigma$ , car  $\bigvee_{i=-n}^n \sigma^{-i} \xi_0$  est la partition mesurable par les cylindres de longueur  $2n+1$  partant en  $-n$ , et puisque ces cylindres lorsque  $n$  parcourt  $\mathbb{N}$ , engendrent la  $\sigma$ -algèbre produit de  $\Sigma$ . Nous avons donc, d'après le théorème de Kolmogorov-Sinai 5.3 (8),

$$h_{\mathbb{P}_{\pi,\nu}}(\sigma) = h_{\mathbb{P}_{\pi,\nu}}(\sigma, \xi_0) = \lim_{n \rightarrow \infty} \frac{1}{2n+1} H_{\mathbb{P}_{\pi,\nu}} \left( \bigvee_{i=-n}^n \sigma^{-i} \xi_0 \right).$$

Or pour tout  $n \in \mathbb{N} \setminus \{0, 1\}$ , en utilisant l'égalité

$$\ln \mathbb{P}_{\pi,\nu}([\alpha_{-n}, \dots, \alpha_n]_{-n}) = \ln \left( \nu_{\alpha_{-n}} \prod_{k=-n}^{n-1} \pi_{\alpha_k, \alpha_{k+1}} \right) = \ln \nu_{\alpha_{-n}} + \sum_{k=-n}^{n-1} \ln \pi_{\alpha_k, \alpha_{k+1}},$$

le fait que  $\nu \pi^p = \nu$  pour tout  $p \in \mathbb{N}$ ,<sup>130</sup> et le fait que la somme des coefficients de chaque ligne de  $\pi^p$  est égale à 1,<sup>131</sup> nous avons

$$\begin{aligned} H_{\mathbb{P}_{\pi,\nu}} \left( \bigvee_{i=-n}^n \sigma^{-i} \xi_0 \right) &= - \sum_{\alpha_{-n}, \dots, \alpha_n \in \mathcal{A}} \mathbb{P}_{\pi,\nu}([\alpha_{-n}, \dots, \alpha_n]_{-n}) \ln \mathbb{P}_{\pi,\nu}([\alpha_{-n}, \dots, \alpha_n]_{-n}) \\ &= - \sum_{\alpha_{-n} \in \mathcal{A}} \nu_{\alpha_{-n}} \ln \nu_{\alpha_{-n}} \sum_{\alpha_{-n+1}, \dots, \alpha_n \in \mathcal{A}} \prod_{i=-n}^{n-1} \pi_{\alpha_i, \alpha_{i+1}} \\ &\quad - \sum_{k=-n}^{n-1} \sum_{\alpha_k, \alpha_{k+1} \in \mathcal{A}} \left( \sum_{\alpha_{-n}, \dots, \alpha_{k-1} \in \mathcal{A}} \nu_{\alpha_{-n}} \prod_{i=-n}^{k-1} \pi_{\alpha_i, \alpha_{i+1}} \right) \pi_{\alpha_k, \alpha_{k+1}} \\ &\quad \quad \quad \times \ln \pi_{\alpha_k, \alpha_{k+1}} \left( \sum_{\alpha_{k+2}, \dots, \alpha_n \in \mathcal{A}} \prod_{i=k+1}^{n-1} \pi_{\alpha_i, \alpha_{i+1}} \right) \\ &= - \sum_{\alpha \in \mathcal{A}} \nu_{\alpha} \ln \nu_{\alpha} - 2n \sum_{\alpha, \beta \in \mathcal{A}} \nu_{\alpha} \pi_{\alpha\beta} \ln(\pi_{\alpha\beta}). \end{aligned}$$

En divisant par  $2n+1$  et en faisant tendre  $n$  vers  $+\infty$ , ceci montre le résultat.  $\square$

Si une matrice de transition  $\pi$  sur l'alphabet  $\mathcal{A}$  est apériodique, donc admet une unique mesure  $\pi$ -stationnaire  $\nu^\pi$  sur  $\mathcal{A}$  par le lemme 6.2 (2), nous noterons  $\mathbb{P}_\pi$  et  $\mathbb{P}_{\pi,+}$  les mesures  $\mathbb{P}_{\pi,\nu^\pi}$  et  $\mathbb{P}_{\pi,\nu^\pi,+}$  respectivement. Le résultat suivant montre que les systèmes de Markov bilatère  $(\Sigma_\pi, \mathbb{P}_\pi, \sigma_\pi)$  et unilatère  $(\Sigma_{\pi,+}, \mathbb{P}_\pi, \sigma_{\pi,+})$  sont mélangeants. L'hypothèse d'apériodicité est nécessaire, comme nous le voyons facilement en prenant pour  $\pi$  la matrice identité.

130. ce qui implique que  $\sum_{\alpha_{-n}, \dots, \alpha_{k-1} \in \mathcal{A}} \nu_{\alpha_{-n}} \prod_{i=-n}^{k-1} \pi_{\alpha_i, \alpha_{i+1}} = \nu_{\alpha_k}$  pour tous les  $k \in \{-n+1, \dots, n\}$  et  $\alpha_k \in \mathcal{A}$

131. c'est-à-dire  $\sum_{\beta_1, \dots, \beta_p \in \mathcal{A}} \prod_{i=0}^{p-1} \pi_{\beta_i, \beta_{i+1}} = 1$  pour tous les éléments  $\beta_0 \in \mathcal{A}$ , ce qui implique que  $\sum_{\alpha_{k+2}, \dots, \alpha_n \in \mathcal{A}} \prod_{i=k+1}^{n-1} \pi_{\alpha_i, \alpha_{i+1}} = 1$  pour tous les  $k \in \{-n-1, \dots, n-2\}$  et  $\alpha_{k+1} \in \mathcal{A}$



**Proposition 6.5.** *Soit  $\pi$  une matrice de transition apériodique sur  $\mathcal{A}$ . Alors le décalage de Bernoulli bilatère  $\sigma$  sur  $\Sigma$  est mélangeant pour la mesure de Markov  $\mathbb{P}_\pi$ , et le décalage de Bernoulli unilatère  $\sigma_+$  sur  $\Sigma_+$  est mélangeant pour la mesure de Markov  $\mathbb{P}_{\pi,+}$ .*

**Démonstration.** Nous ne considérons que le cas bilatère, le cas unilatère se traite de même.

Reprenons les notations de la démonstration du lemme 6.3. Puisque les cylindres engendrent la tribu borélienne de  $\Sigma$ , le sous-espace vectoriel de  $\mathbb{L}^2(\Sigma, \mathbb{P}_\pi)$  engendré par les fonctions caractéristiques  $\mathbb{1}_C$  des cylindres  $C$  est dense dans  $\mathbb{L}^2(\Sigma, \mathbb{P}_\pi)$ . D'après la proposition 3.1 (2), il suffit donc de montrer que, pour tous les cylindres  $C_1, C_2$  de  $\Sigma$ , nous avons

$$\lim_{N \rightarrow \infty} \mathbb{P}_\pi(C_1 \cap \sigma^{-N} C_2) = \mathbb{P}_\pi(C_1) \mathbb{P}_\pi(C_2).$$

Nous pouvons supposer que  $C_1 = [\alpha_m, \dots, \alpha_n]_m$  et  $C_2 = [\beta_p, \dots, \beta_q]_p$  où  $m \leq n$  et  $p \leq q$ . Nous avons alors, pour  $N > n - p$ ,

$$\mathbb{P}_\pi(C_1 \cap \sigma^{-N} C_2) = \nu_{\alpha_m}^\pi \left( \prod_{i=m}^{n-1} \pi_{\alpha_i \alpha_{i+1}} \right) (\pi^{N-n+p})_{\alpha_n \beta_p} \left( \prod_{i=p}^{q-1} \pi_{\beta_i \beta_{i+1}} \right).$$

Pour conclure, puisque  $\mathbb{P}_\pi(C_1) = \nu_{\alpha_m}^\pi \left( \prod_{i=m}^{n-1} \pi_{\alpha_i \alpha_{i+1}} \right)$  et  $\mathbb{P}_\pi(C_2) = \nu_{\beta_p}^\pi \left( \prod_{i=p}^{q-1} \pi_{\beta_i \beta_{i+1}} \right)$ , il suffit donc de montrer que, pour tous les  $\alpha, \beta$  dans  $\mathcal{A}$ , nous avons

$$\lim_{N \rightarrow \infty} (\pi^N)_{\alpha\beta} = \nu_\beta^\pi.$$

Or, par le lemme 6.2, le polynôme caractéristique  $\chi(X) = \det(X \text{id} - {}^t\pi)$  de la matrice  ${}^t\pi$  s'écrit  $\chi(X) = (X - 1)Q(X)$  où  $Q$  est un polynôme premier avec  $X - 1$ , dont les racines complexes sont les valeurs propres complexes de  ${}^t\pi$  différentes de 1, qui sont donc toutes de module strictement inférieur à 1. Donc par le lemme des noyaux, l'espace vectoriel  $\mathbb{R}^{\mathcal{A}}$  est une somme directe  $\mathbb{R}^{\mathcal{A}} = \mathbb{R}\nu^\pi \oplus H_0$  où  $\mathbb{R}\nu^\pi = \ker({}^t\pi - \text{id})$  est la droite vectorielle propre de valeur propre 1 pour  ${}^t\pi$ , et  $H_0 = \ker Q({}^t\pi)$  est un hyperplan vectoriel stable par  ${}^t\pi$  tel que le rayon spectral de la restriction  ${}^t\pi|_{H_0}$  soit strictement inférieur à 1. En particulier  $({}^t\pi|_{H_0})^n$  tend vers 0 dans  $\mathcal{L}(H)$  quand  $n \rightarrow +\infty$ . Par conséquent, la limite  $\tau = \lim_{N \rightarrow \infty} ({}^t\pi)^N$  est le projecteur sur la droite vectorielle  $\mathbb{R}\nu^\pi$  parallèlement à l'hyperplan vectoriel  $H_0$ . De plus, par passage à la limite, la somme des coefficients de chaque colonne de  $\tau$  est égale à 1. Les coefficients matriciels de  $\tau$  sont donc  $\tau_{\alpha\beta} = \nu_\alpha^\pi$  pour tous les  $\alpha, \beta \in \mathcal{A}$ , et  $H_0 = \{v \in \mathbb{R}^{\mathcal{A}} : \sum_{\alpha \in \mathcal{A}} v_\alpha = 0\}$ . Par conséquent  $\lim_{N \rightarrow \infty} (\pi^N)_{\alpha\beta} = {}^t\tau_{\alpha\beta} = \tau_{\beta\alpha} = \nu_\beta^\pi$  pour tous les  $\alpha, \beta \in \mathcal{A}$ , comme voulu.  $\square$

**Remarque.** Avec l'exercice E.39, ceci montre en particulier qu'un système dynamique symbolique  $(\mathcal{A}^{\mathbb{Z}}, \sigma)$  sur un alphabet fini  $\mathcal{A}$  ayant au moins deux éléments admet une famille non dénombrable, à plusieurs paramètres réels, de mesures de probabilité invariantes, mélangeantes et deux à deux étrangères.

## 6.2 Théorie ergodique de la transformation de Gauss <sup>126</sup>

### 6.2.1 Fractions continues

Rappelons que si  $u$  est un nombre réel, alors nous notons  $[u] = \max\{n \in \mathbb{Z} : n \leq u\}$  la partie entière inférieure de  $u$  et  $\{u\} = u - [u]$  sa *partie fractionnaire*.

Il est facile d'approcher tout réel  $x$  par un rationnel  $\frac{p}{q}$  de dénominateur  $q = N$  avec une erreur d'au plus  $\frac{1}{2N}$ . Mieux, si nous autorisons un dénominateur  $q \leq N$ , nous pouvons l'approcher avec une erreur d'au plus  $\frac{1}{q^2}$  (c'est un *théorème de Dirichlet*). Cela résulte du lemme des tiroirs : découpons l'intervalle  $[0, 1]$  en  $N$  intervalles  $I_k = [\frac{k}{N}, \frac{k+1}{N}]$  ; parmi les  $N + 1$  valeurs de  $\{nx\}$  pour  $0 \leq n \leq N$ , deux d'entre elles  $\{n_1x\}$  et  $\{n_2x\}$  sont dans le même intervalle et le réel  $(n_2 - n_1)x$  diffère d'un entier  $p$  par au plus  $\frac{1}{N}$ . Donc (en supposant  $n_2 > n_1$ ) avec  $q = n_2 - n_1$ , nous avons

$$\left| x - \frac{p}{q} \right| \leq \frac{1}{qN} \leq \frac{1}{q^2} .$$

Le développement en fraction continue est une méthode qui permet de trouver rapidement une telle approximation rationnelle. Voici cette méthode.

Pour  $x$  dans  $]0, 1]$ , notons  $a_1 = a_1(x) = \lfloor \frac{1}{x} \rfloor$ , et

$$\psi(x) = \left\{ \frac{1}{x} \right\} = \frac{1}{x} - \lfloor \frac{1}{x} \rfloor = \frac{1}{x} - a_1(x) .$$

Pour  $n \geq 1$ , posons, lorsque c'est défini,  $a_n = a_n(x) = a_1(\psi^{n-1}x)$ . Remarquons que  $a_n \geq 1$  car  $\frac{1}{x} \geq 1$  si  $x \in ]0, 1]$ . Comme  $y = \frac{1}{a_1(y) + \psi(y)}$  pour tout  $y$  dans  $]0, 1]$ , nous avons par récurrence l'égalité

$$x = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_{n-1} + \frac{1}{a_n + \psi^n(x)}}}}} .$$

Posons  $x_n = \psi^n(x) \in [0, 1]$  pour tout  $n \in \mathbb{N}$ , de sorte que

$$x_n = \psi^n(x) = \frac{1}{a_1(\psi^n(x)) + \psi^{n+1}(x)} = \frac{1}{a_{n+1} + x_{n+1}} .$$

Posons  $p_0 = 0$ ,  $q_0 = 1$ ,  $p_1 = 1$  et  $q_1 = a_1$  et par récurrence, pour tout  $n \geq 1$ ,

$$\begin{cases} p_{n+1} = a_{n+1} p_n + p_{n-1} \\ q_{n+1} = a_{n+1} q_n + q_{n-1} \end{cases} . \quad (54)$$

Nous pouvons vérifier par récurrence que

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1} . \quad (55)$$

En particulier,  $p_n$  et  $q_n$  sont premiers entre eux. Par récurrence, nous obtenons que

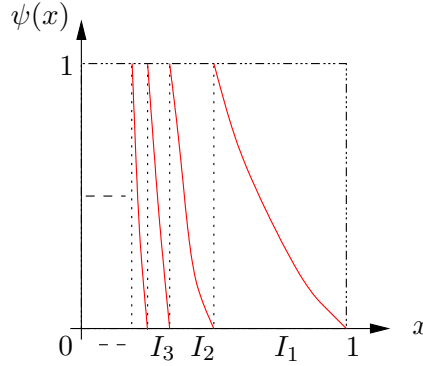
$$x = \frac{p_n + p_{n-1} x_n}{q_n + q_{n-1} x_n} .$$

Tout ceci est bien défini tant que  $\psi^n(x)$  ne s'annule pas, et le *développement en fraction continue*  $x \mapsto (a_1(x), a_2(x), \dots)$  s'arrête si et seulement si  $x$  est rationnel.

Soit  $X$  l'espace topologique  $[0, 1] - \mathbb{Q}$  muni de la topologie induite de  $\mathbb{R}$ . L'application  $\psi : X \rightarrow X$ , définie par  $\psi(x) = \{\frac{1}{x}\}$ , est appelée la *transformation de Gauss*. Pour tout nombre irrationnel  $x \in \mathbb{R} \setminus \mathbb{Q}$ , nous avons, en posant  $a_n = a_n(x - [x])$ ,

$$x = [x] + \lim_{n \rightarrow +\infty} \frac{1}{a_1 + \frac{1}{\dots + \frac{1}{a_{n-1} + \frac{1}{a_n}}}}$$

L'application  $\psi$  est continue à gauche sur  $]0, 1]$ , et discontinue en tout  $\frac{1}{k}$  pour  $k \in \mathbb{N} \setminus \{0, 1\}$ . Elle préserve  $[0, 1] \setminus \mathbb{Q}$  et est continue sur  $[0, 1] \setminus \mathbb{Q}$ . Le graphe de la transformation de Gauss  $\psi$  sur son domaine de définition dans  $[0, 1]$  est le suivant.



**Lemme 6.6.** *Considérons l'alphabet  $\mathcal{A} = \mathbb{N} \setminus \{0\}$ . L'application  $\Theta : x \mapsto (a_1(x), a_2(x), \dots)$  est un homéomorphisme de l'espace  $[0, 1] \setminus \mathbb{Q}$ , muni de la topologie induite de  $\mathbb{R}$ , sur l'espace  $\mathcal{A}^{\mathbb{N} \setminus \{0\}}$ , muni de la topologie produit de la topologie discrète sur  $\mathcal{A}$ . De plus, si  $\sigma_+$  est le décalage à gauche sur  $\mathcal{A}^{\mathbb{N} \setminus \{0\}}$ , lors le diagramme suivant commute :*

$$\begin{array}{ccc} [0, 1] \setminus \mathbb{Q} & \xrightarrow{\Theta} & \mathcal{A}^{\mathbb{N} \setminus \{0\}} \\ \psi \downarrow & & \downarrow \sigma_+ \\ [0, 1] \setminus \mathbb{Q} & \xrightarrow{\Theta} & \mathcal{A}^{\mathbb{N} \setminus \{0\}} \end{array}$$

La dernière affirmation dit que les systèmes dynamiques topologiques  $([0, 1] \setminus \mathbb{Q}, \psi)$  et  $(\mathcal{A}^{\mathbb{N} \setminus \{0\}}, \sigma_+)$  sont topologiquement conjugués. Notons que l'alphabet  $\mathcal{A} = \mathbb{N} \setminus \{0\}$  est infini.

**Démonstration.** Nous ne vérifierons que l'injectivité, par des arguments qui nous réserveront plus tard. Le reste des affirmations est laissé en exercice.

Notons  $I_\ell = a_1^{-1}(\ell) = ]\frac{1}{\ell+1}, \frac{1}{\ell}]$  pour tout  $\ell$  dans  $\mathcal{A}$ , et, pour  $a = (a_1, \dots, a_n) \in \mathcal{A}^n$ , notons

$$\begin{aligned} I_a &= I_{a_1} \cap \psi^{-1}(I_{a_2}) \cap \dots \cap \psi^{-n+1}(I_{a_n}) \\ &= \{x \in [0, 1] \setminus \mathbb{Q} : a_1(x) = a_1, \dots, a_n(x) = a_n\} = \Theta^{-1}([a_1, \dots, a_n]). \end{aligned}$$

D'après les calculs précédents, l'application  $\psi^n$  induit une bijection de  $I_a$  sur  $[0, 1[$ , dont l'inverse est monotone et donné par la formule

$$y \mapsto \frac{p_n + p_{n-1}y}{q_n + q_{n-1}y}$$

où pour tout  $n \in \mathbb{N}$  les entiers positifs ou nuls  $p_n, q_n$  sont définis par les formules (54). Nous pouvons vérifier que  $q_n$  tend en croissant vers  $+\infty$ . Nous en déduisons que  $I_a$  est un intervalle, qui a pour extrémités  $\frac{p_n}{q_n}$  et  $\frac{p_n + p_{n-1}}{q_n + q_{n-1}}$ , et donc que la longueur  $\lambda(I_a)$  de  $I_a$  vérifie, par la formule (55),

$$\lambda(I_a) = \frac{1}{q_n(q_n + q_{n-1})} \leq \frac{1}{q_n^2} \xrightarrow{n \rightarrow \infty} 0.$$

D'où le résultat : pour tout  $x \in [0, 1] \setminus \mathbb{Q}$ , nous avons  $\bigcap_{n \in \mathbb{N}} I_{(a_1(x), \dots, a_n(x))} = \{x\}$ .  $\square$

### 6.2.2 Ergodicité de la transformation de Gauss

L'étude des fractions continues est intimement liée à celle du système dynamique mesurable  $(X = [0, 1] \setminus \mathbb{Q}, \psi : x \mapsto \{\frac{1}{x}\})$ .

**Proposition 6.7.**

(1) (**Théorème de Gauss**) *La transformation de Gauss  $\Psi$  préserve la mesure de probabilité  $\mu$  (dite mesure de Gauss) sur  $X$  définie par*

$$\mu = \frac{1}{\ln 2} \left( \frac{dx}{1+x} \right).$$

(2) *La transformation de Gauss  $\Psi$  est ergodique pour la mesure de Gauss  $\mu$ .*

Remarquons que la mesure de Lebesgue et la mesure de Gauss sur  $[0, 1]$  sont absolument continues l'une par rapport à l'autre, c'est-à-dire qu'elles ont les mêmes ensembles de mesure nulle.

**Démonstration.** (1) Pour toute fonction positive mesurable  $f$  sur  $[0, 1]$ , puisque la restriction  $\psi|_{[\frac{1}{n+1}, \frac{1}{n}] \cap X}$  est l'application définie par  $t \mapsto \frac{1}{t} - \lfloor \frac{1}{t} \rfloor = \frac{1}{t} - n$  et par changement de variable, nous avons

$$\begin{aligned} \int_0^1 f(\psi(t)) \frac{dt}{1+t} &= \sum_{n=1}^{\infty} \int_{\frac{1}{n+1}}^{\frac{1}{n}} f\left(\frac{1}{t} - n\right) \frac{dt}{1+t} = \sum_{n=1}^{\infty} \int_0^1 f(s) \frac{ds}{(s+n)^2(1+\frac{1}{s+n})} \\ &= \int_0^1 f(s) \sum_{n=1}^{\infty} \frac{1}{(s+n)(s+n+1)} ds = \int_0^1 f(s) \frac{ds}{1+s}. \end{aligned}$$

(2) Afin de démontrer l'ergodicité de la transformation de Gauss, nous aurons besoin de l'inégalité de gauche dans le lemme suivant (l'inégalité de droite sera utile pour montrer le mélange). Reprenons les notations du lemme 6.6.

**Lemme 6.8. (Condition de Renyi <sup>132</sup>)** Il existe des constantes  $c, c' > 0$  telles que, pour tout borélien  $A$  de  $[0, 1]$  et tout  $a = (a_1, \dots, a_n)$  dans  $\mathcal{A}^n$ , nous avons

$$c \mu(I_a) \mu(A) \leq \mu(I_a \cap \psi^{-n}(A)) \leq c' \mu(I_a) \mu(A).$$

**Démonstration.** Il suffit de montrer cette double inégalité avec la mesure de Lebesgue  $\lambda = dx$  au lieu de la mesure de Gauss  $\mu$ , car nous avons  $\lambda \leq (2 \ln 2) \mu \leq 2\lambda$ . Rappelons que  $\psi^n$  est une bijection de  $I_a$  sur  $[0, 1[$ , d'inverse l'homographie  $\psi_a : y \mapsto \frac{p_n + p_{n-1}y}{q_n + q_{n-1}y}$ . Écrivons alors

$$\lambda(I_a \cap \psi^{-n}(A)) = \int_{\psi_a(A)} dx = \int_A |\psi'_a(y)| dy.$$

Comme  $|\psi'_a(y)| = \frac{1}{(q_n + q_{n-1}y)^2}$  par la formule (55), nous avons  $\lambda(I_a) = \frac{1}{q_n(q_n + q_{n-1})}$ . Puisque  $q_{n-1} \leq q_n$ , nous avons

$$\frac{1}{2} \lambda(I_a) \leq |\psi'_a(y)| \leq 2 \lambda(I_a).$$

D'où le résultat, par intégration sur  $y \in A$ .  $\square$

**Fin de la démonstration de la proposition 6.7.** Soit  $A$  un borélien de  $X$  tel que  $\psi^{-1}(A) = A$ . D'après la condition de Renyi 6.8, la mesure borélienne sur l'espace mesurable  $X$  définie par  $Y \mapsto \nu(Y) = \mu(A \cap Y) - c \mu(A) \mu(Y)$  prend des valeurs positives pour tout  $Y = I_a$ . Elle prend donc, par approximation, des valeurs positives sur toutes les fonctions continues positives. Donc  $\nu$  est une mesure positive. Nous avons donc  $\nu({}^c A) = -c \mu(A) \mu({}^c A) \geq 0$ . Donc  $\mu(A)$  vaut 0 ou 1, et la mesure de Gauss  $\mu$  est ergodique.  $\square$

**Exercice E.37.** Avec les notations du lemme 6.6, pour tout cylindre initial  $[a_1, \dots, a_n]$  de  $\mathcal{A}^{\mathbb{N} \setminus \{0\}}$ , calculer sa mesure pour la loi image  $\Theta_* \mu$  en fonction des approximations rationnelles  $\frac{p_n}{q_n}$ .

Nous en déduisons la propriété suivante sur la fréquence d'apparition d'un entier  $\ell$  dans le développement en fractions continues d'un réel  $x$ .

**Corollaire 6.9.** Pour Lebesgue-presque tout  $x$  dans  $[0, 1]$ , pour tout  $\ell$  dans  $\mathbb{N} - \{0\}$ , nous avons

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{Card} \{k \leq n : a_k(x) = \ell\} = \log_2 \left( 1 + \frac{1}{\ell(\ell + 2)} \right).$$

**Démonstration.** Le théorème ergodique de Birkhoff, appliqué à la transformation de Gauss  $\psi$  (qui est ergodique par la proposition 6.7 (2)) et à la fonction caractéristique (donc  $\mu$ -intégrable)  $f = \mathbb{1}_{\{a_1^{-1}(\ell)\}}$ , affirme que, pour  $\mu$ -presque tout  $x \in [0, 1]$  (donc Lebesgue-presque tout  $x \in [0, 1]$ , cette limite existe et est égale à

$$\begin{aligned} \int f d\mu &= \frac{1}{\ln 2} \int_{\frac{1}{\ell+1}}^{\frac{1}{\ell}} \frac{dx}{1+x} = \frac{1}{\ln 2} \left( \ln \left( 1 + \frac{1}{\ell} \right) - \ln \left( 1 + \frac{1}{\ell+1} \right) \right) \\ &= \log_2 \left( 1 + \frac{1}{\ell(\ell+2)} \right). \quad \square \end{aligned}$$



Alfred Rényi  
1921-1970



Émile Borel  
1871-1956



Felix Bernstein  
1878-1956

### 6.2.3 Mélange de la transformation de Gauss

Avec un peu plus de travail, nous pouvons préciser la proposition 6.7.

**Proposition 6.10.** *La transformation de Gauss  $\psi$  est mélangeante pour la mesure de Gauss  $\mu$ .*

Le lemme suivant nous donne un renseignement très proche du mélange.

**Lemme 6.11.** *Il existe des constantes  $c, c' > 0$  telles que, pour tous les boréliens  $A, B$  de  $[0, 1]$ , nous ayons*

$$c \mu(A)\mu(B) \leq \liminf_{n \rightarrow \infty} \mu(A \cap \psi^{-n}(B)) \leq \limsup_{n \rightarrow \infty} \mu(A \cap \psi^{-n}(B)) \leq c' \mu(A)\mu(B). \quad (56)$$

**Démonstration.** Reprenons de nouveau les notations du lemme 6.6. Notons  $\mathcal{B}_n$  la  $\sigma$ -algèbre sur  $X$  engendrée par les  $I_a \cap X$  avec  $a = (a_1, \dots, a_n) \in \mathcal{A}$  et  $n \in \mathbb{N} \setminus \{0\}$ . D'après la condition de Renyi 6.8, la formule (56) est vraie pour  $A = I_a$ , donc aussi pour  $A$  dans la famille  $\mathcal{B}' = \cup_{n \geq 1} \mathcal{B}_n$ . Par un argument d'approximation, donné ci-dessous, cette assertion est donc vraie pour tout borélien.

Notons  $A \Delta A' = (A - A') \cup (A' - A)$  la différence symétrique de deux parties  $A, A'$ . Comme la  $\sigma$ -algèbre  $\mathcal{B}$  des boréliens de  $[0, 1]$  est engendrée par l'algèbre de Boole  $\mathcal{B}'$ , pour tout  $A$  dans  $\mathcal{B}$ , nous avons

$$\forall \epsilon > 0, \quad \exists A' \in \mathcal{B}', \quad \mu(A \Delta A') \leq \epsilon.$$

(En effet, l'ensemble des boréliens  $A$  qui vérifient cette propriété est une  $\sigma$ -algèbre.) L'assertion (56) pour  $A'$  permet de montrer l'assertion pour  $A$  à  $\epsilon$  près. Comme  $\epsilon$  est arbitraire, l'assertion (56) pour  $A$  s'en déduit.  $\square$

**Démonstration de la proposition 6.10.** Posons  $Y = [0, 1] \times [0, 1]$ ,  $\phi : Y \rightarrow Y$  définie (presque partout) par  $(x, x') \mapsto (\psi(x), \psi(x'))$ ,  $\nu = \mu \times \mu$  la mesure produit sur  $Y$ , et  $\nu_n = \mu \times (\psi^n)_* \mu$ . Dire que  $\psi$  est mélangeante, c'est dire que  $\nu_n$  converge \*-faiblement vers  $\nu$  (voir la proposition 3.1 (3)). Par la compacité faible de l'espace  $\text{Prob}(Y)$  des mesures de probabilité sur  $Y$ , il suffit de montrer que si une sous-suite  $(\nu_{n_k})_{k \in \mathbb{N}}$  converge \*-faiblement vers une mesure de probabilité  $\nu_\infty$ , alors  $\nu_\infty = \nu$ .

**Affirmation 1.** Pour tous boréliens  $P, Q$  de  $Y$ , nous avons

$$c^2 \nu(P)\nu(Q) \leq \liminf_{n \rightarrow \infty} \nu(P \cap g^{-n}(Q)).$$

**Démonstration.** Cette affirmation résulte du lemme 6.11, lorsque  $P$  et  $Q$  sont des pavés  $P = A_1 \times A_2$  et  $Q = B_1 \times B_2$ , et donc aussi lorsque  $P$  et  $Q$  sont des réunions finies disjointes de pavés. Le cas général s'en déduit par approximation, comme ci-dessus.  $\square$

**Affirmation 2.** La mesure  $\nu$  est ergodique pour  $\phi$ .

**Démonstration.** Si  $Q$  est un borélien de  $Y$  tel que  $\phi^{-1}(Q) = Q$ , alors en prenant  $P = {}^c Q$  dans l'affirmation précédente, nous obtenons que  $\nu(Q)\nu({}^c Q) = 0$ .  $\square$

**Affirmation 3.** Pour tout borélien  $P$  de  $Y$ , nous avons  $\nu_\infty(P) \leq c' \nu(P)$ .

**Démonstration.** Ceci résulte du lemme 6.11 lorsque  $P$  est un pavé  $P = A \times B$  avec  $A$  et  $B$  fermés, car  $\nu_\infty(P) \leq \limsup_{n \rightarrow +\infty} \nu_n(P) \leq c' \nu(P)$ . Le cas général s'en déduit par approximation, comme ci-dessus.  $\square$

Terminons la démonstration de la proposition 6.10. Les mesures  $\nu_\infty$  et  $\nu$  sont des mesures de probabilité  $\phi$ -invariantes sur  $Y$ . La mesure  $\nu_\infty$  est absolument continue par rapport à  $\nu$  par le lemme précédent. Par le théorème de Radon-Nikodym, il existe  $f$  dans  $\mathbb{L}^1(Y, \nu)$  telle que  $\nu_\infty = f\nu$ . La fonction  $f$  est ( $\nu$ -presque partout)  $\phi$ -invariante. Comme  $\nu$  est ergodique, l'application  $f$  est constante ( $\nu$ -presque partout), et  $\nu_\infty = \nu$ .  $\square$

### 6.3 Exercices

**Exercice E.38.** Soit  $A = (A_{\alpha,\beta})_{\alpha,\beta \in \mathcal{A}}$  une matrice carrée de coefficients égaux à 0 ou 1. Nous appellerons *période*  $d$  de  $A$  le plus grand diviseur commun de l'ensemble  $\{n > 0 : \exists \alpha \in \Lambda, (A^n)_{\alpha\alpha} \neq 0\}$ , avec la convention que  $d = \infty$  si cet ensemble est vide. Montrer que  $A$  est apériodique si et seulement si  $A$  est irréductible de période 1.

**Exercice E.39.** Soient  $\pi$  et  $\pi'$  deux matrices de transition apériodiques distinctes, indexées par un même alphabet fini  $\mathcal{A}$ . Montrer que les mesures de Markov  $\mathbb{P}_\pi$  et  $\mathbb{P}_{\pi'}$  sur  $\mathcal{A}^{\mathbb{Z}}$  sont étrangères.

**Exercice E.40.** Soient  $\mathcal{A}$  un alphabet fini,  $\pi$  une matrice de transition sur  $\mathcal{A}$ ,  $\nu$  une mesure  $\pi$ -stationnaire sur  $\mathcal{A}$ , et  $\mathbb{P} = \mathbb{P}_{\pi,\nu}$  ou  $\mathbb{P} = \mathbb{P}_{\pi,\nu,+}$  la mesure de Markov bilatère ou unilatère associée à  $(\pi, \nu)$ . Montrer que si  $\pi$  est irréductible, alors la mesure  $\pi$ -stationnaire  $\nu$  est de support total dans  $\mathcal{A}$  et, en étudiant la convergence en moyenne des puissances de  $\pi$ , qu'elle est unique. Montrer que les propriétés suivantes sont équivalentes.

- (1) Les systèmes de Markov bilatères et unilatères associés à  $(\pi, \nu)$  sont ergodiques et la mesure  $\pi$ -stationnaire  $\nu$  est de support total sur  $\mathcal{A}$ .
- (2) La matrice de transition  $\pi$  est irréductible.
- (3) La mesure  $\pi$ -stationnaire  $\nu$  est de support total sur  $\mathcal{A}$ .

Pour monter que (2) implique (1), on pourra utiliser la caractérisation de l'ergodicité par la convergence en moyenne de Césaro des coefficients de corrélations sur un  $\pi$ -système (voir l'exercice E.9 (6)).

Donner un exemple où le système de Markov bilatère associé à  $(\pi, \nu)$  est ergodique, mais où la mesure  $\pi$ -stationnaire  $\nu$  n'est de support total sur  $\mathcal{A}$ .

**Exercice E.41.** Soit  $\mathcal{A}$  l'alphabet  $\{a, 1, 2, b\}$ .

- (1) Montrer qu'il existe une matrice de transition  $\pi = (\pi_{\alpha,\beta})_{\alpha,\beta \in \mathcal{A}}$  sur  $\mathcal{A}$  dont les coefficients non nuls sont exactement  $\pi_{a,1}, \pi_{1,1}, \pi_{1,b}, \pi_{2,a}, \pi_{b,2}$ , et qu'elle admet une unique mesure  $\pi$ -stationnaire  $\nu$ . Notons  $\mathcal{B}$  la  $\sigma$ -algèbre produit de  $\mathcal{A}^{\mathbb{Z}}$ ,  $\mu = \mathbb{P}_\pi$  la mesure de Markov bilatère associée sur  $(\mathcal{A}^{\mathbb{Z}}, \mathcal{B})$ ,  $\Sigma$  son support et  $\phi$  la restriction à  $\Sigma$  du décalage à gauche sur  $\mathcal{A}^{\mathbb{Z}}$ .
- (2) Montrer que le système dynamique  $(\Sigma, \mathcal{B}, \mu, \phi)$  est mélangeant.
- (3) Notons  $A = \{(x_n)_{n \in \mathbb{N}} \in \Sigma : x_0 \in \{a, b\}\}$ . Montrer que le système dynamique induit  $(A, \mathcal{B}_A, \mu_A, \phi_A)$  (voir l'exercice E.13) n'est pas mélangeant.

### 6.4 Indications pour la résolution des exercices

**Exercice E.40.** Nous commençons la démonstration par des remarques préliminaires.

Pour tous les cylindres centrés  $C = [\alpha_{-k}, \alpha_k]_{-k}$  et  $D = [\beta_{-\ell}, \beta_\ell]_{-\ell}$  dans  $\mathcal{A}^{\mathbb{Z}}$ , et par la définition de la mesure de Markov bilatère  $\mathbb{P} = \mathbb{P}_{\pi, \nu}$ , pour tout  $n \geq k + \ell$ , nous avons

$$\mathbb{P}(C \cap \sigma^{-n}D) = \nu_{\alpha_{-k}} \left( \prod_{i=-k}^{k-1} \pi_{\alpha_i \alpha_{i+1}} \right) (\pi^{n-k-\ell})_{\alpha_k \beta_{-\ell}} \left( \prod_{j=-\ell}^{\ell-1} \pi_{\beta_j \beta_{j+1}} \right). \quad (57)$$

En appliquant cette formule au cas particulier où  $k = \ell = 0$ ,  $C = [\alpha]$  et  $D = \beta$ , nous avons  $\mathbb{P}([\alpha] \cap \sigma^{-n}[\beta]) = \nu_\alpha (\pi^n)_{\alpha\beta}$ .

Par le théorème de Birkhoff 2.2 appliqué à la fonction caractéristique  $\mathbb{1}_{[\beta]}$ , qui est  $\mathbb{L}^1$ , la moyenne de Birkhoff  $S_n \mathbb{1}_{[\beta]} = \frac{1}{n} \sum_{k=1}^{n-1} \mathbb{1}_{[\beta]} \circ \sigma^k$  converge dans  $\mathbb{L}^1$ , donc faiblement dans  $\mathbb{L}^2$  par l'inégalité de Cauchy-Schwarz, vers une fonction intégrable  $\widetilde{\mathbb{1}_{[\beta]}}$ . Donc pour tout  $\alpha \in \mathcal{A}$  tel que  $\nu_\alpha > 0$ , nous avons donc

$$\begin{aligned} \nu_\alpha \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^{n-1} (\pi^k)_{\alpha\beta} &= \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^{n-1} \mathbb{P}([\alpha] \cap \sigma^{-k}[\beta]) \\ &= \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^{n-1} \int_{\mathcal{A}^{\mathbb{Z}}} \mathbb{1}_{[\alpha]} \mathbb{1}_{[\beta]} \circ \sigma^k d\mathbb{P} \\ &= \lim_{n \rightarrow +\infty} \int_{\mathcal{A}^{\mathbb{Z}}} \mathbb{1}_{[\alpha]} \left( \frac{1}{n} \sum_{k=1}^{n-1} \mathbb{1}_{[\beta]} \circ \sigma^k \right) d\mathbb{P} = \int_{\mathcal{A}^{\mathbb{Z}}} \mathbb{1}_{[\alpha]} \widetilde{\mathbb{1}_{[\beta]}} d\mathbb{P}. \end{aligned} \quad (58)$$

Montrons tout d'abord que l'assertion (2) implique l'assertion (3). Soit  $\beta \in \mathcal{A}$ . Pour tout  $\gamma \in \mathcal{A}$  tel que  $\nu_\gamma > 0$  (qui existe car  $\nu$  est une mesure de probabilité sur  $\mathcal{A}$ ), puisque  $\pi$  est irréductible, il existe  $n \geq 1$  tel que  $(\pi^n)_{\gamma\beta} > 0$ . Or  $\nu = \nu \pi^k$  puisque  $\nu$  est  $\pi$ -stationnaire. Donc

$$\nu_\beta = \sum_{\alpha \in \mathcal{A}} \nu_\alpha (\pi^n)_{\alpha\beta} \geq \nu_\gamma (\pi^n)_{\gamma\beta} > 0,$$

ce qui montre l'assertion (3).

Montrons que l'assertion (3) implique l'unicité de  $\nu$  et l'assertion (1) dans le cas bilatère, le résultat dans le cas unilatère est analogue. Supposons donc que  $\nu_\alpha > 0$  pour tout  $\alpha \in \mathcal{A}$ . La formule (58) montre que la limite  $\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^{n-1} (\pi^k)_{\alpha\beta}$  existe. Notons-la  $Q = (q_{\alpha\beta})_{\alpha, \beta \in \mathcal{A}}$ . C'est une matrice de transition (comme vu dans la démonstration du lemme 6.2 (1)), et de plus par l'effet de la moyenne, nous avons  $Q\pi = \pi Q = Q$  et  $Q^2 = Q$ .

Montrons que les lignes de la matrice  $Q$  sont égales. Pour tout  $\beta \in \mathcal{A}$ , soit  $\alpha \in \mathcal{A}$  tel que  $q_{\alpha\beta} = \max_{\alpha' \in \mathcal{A}} q_{\alpha'\beta}$ . Puisque  $Q^2 = Q$ , nous avons  $q_{\alpha\beta} = \sum_{\gamma \in \mathcal{A}} q_{\alpha\gamma} q_{\gamma\beta}$ , donc  $q_{\alpha\beta}$  est une moyenne (puisque la somme des coefficients de la ligne d'indice  $\alpha$  de  $Q$  est égale à 1) de nombres positifs  $q_{\gamma\beta}$  pour  $\gamma \in \mathcal{A}$ , qui sont tous inférieurs ou égaux à  $q_{\alpha\beta}$ . Les nombres  $q_{\gamma\beta}$  pour  $\gamma \in \mathcal{A}$  sont donc tous égaux, ce qui montre le résultat.

Montrons l'unicité de la mesure  $\pi$ -stationnaire  $\nu$ . Soit  $\nu'$  une mesure  $\pi$ -stationnaire. Puisque  $\nu' \pi = \nu'$ , en moyennant cette formule, en passant à la limite et en utilisant le fait que les coefficients de toute colonne fixée de  $Q$  sont égaux et que la somme des coefficients de  $\nu$  est égale à 1, nous avons donc

$$\nu' = \nu' \frac{1}{n} \sum_{k=1}^{n-1} \pi^k = \nu' Q = \nu,$$



ce qui montre le résultat. En particulier, comme les lignes de  $Q$  sont des mesures  $\pi$ -stationnaires, pour tous les  $\alpha, \beta \in \mathcal{A}$ , nous avons  $\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^{n-1} (\pi^k)_{\alpha\beta} = q_{\alpha\beta} = \nu_\beta$ . Par la formule (57), nous avons donc

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^{n-1} \mathbb{P}(C \cap \sigma^{-n}D) = \nu_{\alpha_{-k}} \left( \prod_{i=-k}^{k-1} \pi_{\alpha_i \alpha_{i+1}} \right) \nu_{\beta_{-\ell}} \left( \prod_{j=-\ell}^{\ell-1} \pi_{\beta_j \beta_{j+1}} \right) = \mathbb{P}(C) \mathbb{P}(D).$$

En utilisant la caractérisation de l'ergodicité par la convergence en moyenne de Césaro des coefficients de corrélations sur le  $\pi$ -système formé des cylindres centrés de  $\mathcal{A}^{\mathbb{Z}}$  (voir l'exercice E.9 (6)), ceci montre l'ergodicité de  $(\mathcal{A}^{\mathbb{Z}}, \mathbb{P}, \sigma)$ .

Montrons enfin que l'assertion (1) dans le cas bilatère implique l'assertion (2), le résultat dans le cas unilatère est analogue. Si  $(\mathcal{A}^{\mathbb{Z}}, \mathbb{P}, \sigma)$  est ergodique, alors avec les notations ci-dessus, la fonction  $\mathbb{1}_{[\beta]}$  est  $\mathbb{P}$ -presque partout constante égale à  $\int_{\mathcal{A}^{\mathbb{Z}}} \mathbb{1}_{[\beta]} d\mathbb{P} = \nu_\beta$  pour tout  $\beta \in \mathcal{A}$ . Donc par la formule (58), pour tous les  $\alpha, \beta \in \mathcal{A}$ , nous avons  $\nu_\alpha \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^{n-1} (\pi^k)_{\alpha\beta} = \nu_\alpha \nu_\beta$ . Comme  $\nu_\alpha$  et  $\nu_\beta$  sont non nuls, ceci implique que  $\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^{n-1} (\pi^k)_{\alpha\beta}$  est non nul. Donc a fortiori il existe  $k \in \mathbb{N}$  tel que  $(\pi^k)_{\alpha\beta} \neq 0$ , ce qui montre l'assertion (2).

Soit  $\pi'$  une matrice de transition irréductible pour un alphabet  $\mathcal{A}'$ , et  $\nu'$  son unique mesure  $\pi'$ -stationnaire. Considérons l'alphabet  $\mathcal{A} = \{\omega\} \cup \mathcal{A}'$  où  $\omega \notin \mathcal{A}'$ , la matrice  $\pi = \begin{pmatrix} 1 & 0 \\ 0 & \pi' \end{pmatrix}$  (de sorte que  $\pi_{\omega, \omega} = 1$  et  $\pi_{\alpha, \beta} = 0$  si  $\alpha = \omega$  ou  $\beta = \omega$ ) et  $\nu = (\nu_\omega = 0, \nu')$ . Alors il est immédiat de vérifier que  $\pi$  est une matrice de transition non irréductible et que  $\nu$  est une mesure  $\pi$ -stationnaire. Soit  $\mathbb{P} = \mathbb{P}_{\pi, \nu}$ . L'ensemble  $(\mathcal{A}')^{\mathbb{Z}}$  est une partie  $\sigma$ -invariante de  $\mathcal{A}^{\mathbb{Z}}$  de  $\mathbb{P}$ -mesure totale, car son complémentaire est  $\bigcup_{n \in \mathbb{Z}} \{(\alpha_i)_{i \in \mathbb{Z}} \in \mathcal{A}^{\mathbb{Z}} : \alpha_n = \omega\}$ , qui est de mesure nulle comme union dénombrable des ensembles  $[\omega]_n$  de mesure nulle  $\mathbb{P}([\omega]_n) = \nu_\omega = 0$ . Puisque  $\pi'$  est irréductible, il découle de l'équivalence précédente que  $(\mathcal{A}^{\mathbb{Z}}, \mathbb{P}, \sigma)$  est ergodique, car la propriété d'ergodicité est invariante par passage à une partie invariante de mesure totale.

**Exercice E.41.** (1) et (2) En utilisant l'ordre  $a, 1, 2, b$  sur l'alphabet  $\mathcal{A}$ , les telles matrices

sont les matrices de la forme  $\pi = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & x & 0 & y \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$  où  $x, y \in ]0, 1[$  vérifient  $x + y = 1$ . Un

calcul élémentaire montre que  $\pi^6 = \begin{pmatrix} x^2y & x^5 + 2xy & x^3y & x^4y + y^2 \\ x^3y & x^6 + 3x^2y & x^4y + y^2 & x^5y + 2xy^2 \\ xy & x^4 + y & x^2y & x^3y \\ y & x^3 & xy & x^2y \end{pmatrix}$ , dont tous

les coefficients sont strictement positifs. Donc la matrice de transition  $\pi$  est apériodique. Il découle alors du lemme 6.2 (2) et de la proposition 6.5 qu'il existe une unique mesure  $\pi$ -stationnaire et que la mesure de Markov associée  $\mu$  est mélangeante pour le décalage bilatère à gauche sur  $\mathcal{A}^{\mathbb{Z}}$ , donc sur le support de  $\mu$ . La mesure  $\pi$ -stationnaire est par ailleurs  $\nu = (\nu_a = \frac{1}{4}, \nu_1 = \frac{1}{4}, \nu_2 = \frac{1}{4}, \nu_b = \frac{1}{4})$  par unicité.

(3) Pour tout  $\alpha \in \mathcal{A}$ , notons  $[\alpha]$  le cylindre  $\{(x_n)_{n \in \mathbb{Z}} \in \Sigma : x_0 = \alpha\}$ . Alors  $A = [a] \sqcup [b]$  et sa mesure est  $\mu(A) = \mu([a]) + \mu([b]) = \nu_a + \nu_b = \frac{1}{2}$ . Pour tout  $k \in \mathbb{N} \setminus \{0\}$ , notons  $1^k = 111 \dots 11$  la suite où le symbole 1 est répété  $k$  fois. La mesure de Markov du cylindre

$[1^k]$  est par définition  $\mu([1^k]) = \nu_1 \pi_{11}^{k-1} = \frac{1}{4x^{k-1}}$ , qui tend vers 0 quand  $k$  tend vers  $+\infty$ . Les éléments de  $[a]$  qui ne commencent pas et ne terminent pas par une infinité de 1 (ceux-ci formant un ensemble de mesure pleine dans  $[a]$ ) sont de la forme

$$x = \dots 1^{n-2} b 2 a 1^{n-1} b 2 a 1^{n_0} | a 1^{n_1} b 2 a 1^{n_2} b 2 a 1^{n_3} b 2 a \dots ,$$

où  $n_i \in \mathbb{N} \setminus \{0\}$  pour tout  $i \in \mathbb{Z}$  et où la barre indique l'endroit après lequel se trouve le coefficient  $x_0$  d'indice 0 de la suite  $x = (x_i)_{i \in \mathbb{Z}}$ . Comme les lettres  $a$  et  $b$  apparaissent de manière alternée dans  $x$ , nous avons donc  $\phi_A^{-2n}([a]) = [a]$  pour tout  $n \in \mathbb{Z}$  (modulo ensemble de mesure nulle), donc  $\mu_A([b] \cap \phi_A^{-2n}([a])) = \mu_A([b] \cap [a]) = 0$  ne converge pas vers  $\mu_A([b])\mu_A([a]) = \frac{\mu([b])\mu([a])}{\mu(A)^2} = \frac{1}{4}$ .

## Deuxième partie

# Systèmes dynamiques topologiques et différentiables

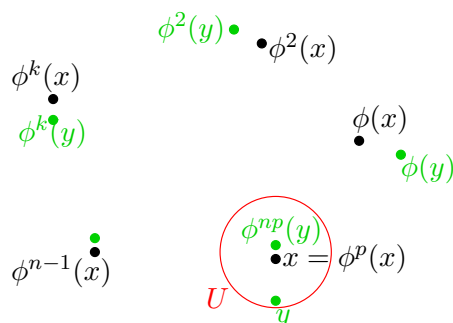
Dans cette partie, nous étudions les systèmes dynamiques topologiques et différentiables, en particulier les comportements asymptotiques de leurs orbites. Mais l'aide des mesures invariantes<sup>133</sup> et donc de la théorie ergodique, est incontournable dans cette étude, et nous utiliserons donc dans cette partie II certains résultats de la partie I, qui peuvent être admis en boîte noire par les personnes n'ayant pas lu cette première partie.

## 7 Généralités de dynamique topologique

### 7.1 Compléments sur les orbites périodiques

Donnons pour commencer, après quelques définitions supplémentaires sur la nature des points périodiques, un exemple de système dynamique topologique supplémentaires par rapport à ceux de la partie 0.

Soient  $(X, \phi)$  un système dynamique topologique à temps discret et  $x \in X$  un point périodique de période  $p$ . Nous dirons que le point périodique  $x$  est (localement) *attractif* s'il existe un voisinage  $U$  de  $x$  tel que pour tout  $y \in U$ , nous avons  $\lim_{n \rightarrow +\infty} \phi^{np}(y) = x$ .



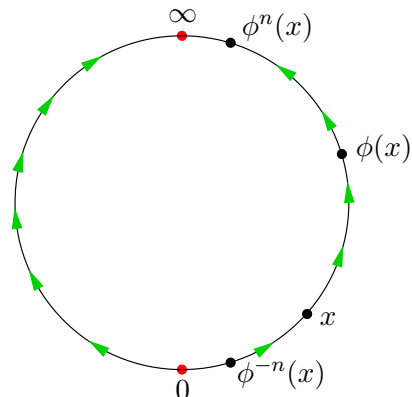
Ceci implique que l'ensemble des points d'accumulation de la suite  $(\phi^n(y))_{n \in \mathbb{N}}$  est exactement l'orbite (finie)  $\mathcal{O}(x)$  de  $x$ . Si le système dynamique  $(X, \phi)$  est inversible, nous dirons que le point périodique  $x$  est *répulsif* s'il est attractif pour le système dynamique  $(X, \phi^{-1})$ . Si  $h$  est une conjugaison entre  $(X, \phi)$  et un système dynamique topologique à temps discret  $(X', \phi')$ , alors  $x$  est un point périodique attractif (respectivement répulsif) de  $(X, \phi)$  si et seulement si  $h(x)$  est un point périodique attractif (respectivement répulsif) de  $(X', \phi')$ .

**Exemple.** Notons  $X = \widehat{\mathbb{R}}$  le compactifié d'Alexandrov<sup>134</sup> de  $\mathbb{R}$ . Pour  $\lambda > 1$  fixé, notons  $\phi : X \rightarrow X$  l'application définie par  $t \mapsto \lambda t$  si  $t \in \mathbb{R}$  et  $\infty \mapsto \infty$ .

133. dont l'existence est assurée par exemple dans le cas des espaces de phases compacts métrisables non vides par le théorème de Krylov-Bogolyubov (voir la proposition 0.6)

134. Si  $Y$  est un espace topologique localement compact, alors le *compactifié d'Alexandrov* de  $Y$  est l'espace topologique compact  $\widehat{Y}$  défini comme l'ensemble  $Y \cup \{\infty\}$  (où  $\infty$  est un ensemble n'appartenant pas à  $Y$ ) muni de la topologie dont les ouverts sont d'une part les ouverts de  $Y$ , d'autre part les parties de la forme  ${}^c K \cup \{\infty\}$  où  $K$  est un compact de  $Y$ . L'espace  $Y$  est ouvert dans  $\widehat{Y}$ , et dense dans  $\widehat{Y}$  s'il n'est pas compact. Voir par exemple l'exercice E.51 de [Pau3].

Ce système a exactement deux points périodiques, qui sont en fait des points fixes, les points 0 et  $\infty$ . Pour tout  $x \in X - \{0, \infty\}$ , nous avons  $\lim_{n \rightarrow +\infty} \phi^n(x) = \infty$  et  $\lim_{n \rightarrow -\infty} \phi^n(x) = 0$ . En particulier, 0 est un point fixe répulsif et  $\infty$  est un point fixe attractif. Voir le dessin ci-contre, où nous rappelons que l'application du cercle  $\mathbb{S}_1$  dans  $\widehat{\mathbb{R}}$ , qui envoie le pôle Nord  $N = (0, 1)$  sur  $\infty$  et qui envoie  $x \in \mathbb{S}_1 - \{N\}$  sur le point d'intersection avec l'axe horizontal  $\mathbb{R}$  de la droite passant par  $N$  et  $x$ , appelée la *projection stéréographique* (de pôle Nord), est un homéomorphisme, qui envoie le pôle Sud  $S = (0, -1)$  sur 0.



Cet exemple de système dynamique topologique est un exemple archétypal de *dynamique Nord-Sud*, qui est un système dynamique topologique  $(X, \phi)$  à temps discret inversible, ou à temps continu, qui admet deux points fixes  $x, y$  tels que pour tout  $z \in X - \{x, y\}$ , nous avons  $\lim_{t \rightarrow +\infty} \phi^t(z) = y$  et  $\lim_{t \rightarrow -\infty} \phi^t(z) = x$ .

Le résultat suivant décrit les périodes des systèmes dynamiques topologiques à temps discret sur un intervalle compact réel.

**Théorème 7.1.** [Théorème de Sharkovsky<sup>135</sup> (1964)] Soient  $I$  un intervalle compact de  $\mathbb{R}$  et  $f : I \rightarrow I$  une application continue.

Si le système dynamique  $(I, f)$  a un point périodique de période 3, alors  $(I, f)$  a un point périodique de période  $k$  pour tout  $k \in \mathbb{N} - \{0\}$ .

Ceci est une version simplifiée, popularisée par Li et Yorke (voir [LY]), du résultat de Sharkovsky. L'ordre de Sharkovsky sur  $\mathbb{N} - \{0\}$  est l'ordre total  $\preceq$  défini par (dans l'ordre décroissant)

$$3 \succ 5 \succ 7 \succ \dots \succ 2 \cdot 3 \succ 2 \cdot 5 \succ 2 \cdot 7 \succ \dots \succ 2^2 \cdot 3 \succ 2^2 \cdot 5 \succ 2^2 \cdot 7 \succ \dots \succ 2^3 \succ 2^2 \succ 2 \succ 1.$$

La version complète du théorème de Sharkovsky dit (voir par exemple [BurH]) qu'une partie de  $\mathbb{N} - \{0\}$  est l'ensemble des périodes d'une application continue de  $[0, 1]$  dans lui-même si et seulement si elle est l'ensemble  $\{\dots, 2^2, 2, 1\}$  des puissances de 2 ou un ensemble  $\{n \in \mathbb{N} - \{0\} : n \preceq n_0\}$  pour un  $n_0 \in \mathbb{N}$ .

**Démonstration.** Le lemme suivant est élémentaire.

**Lemme 7.2.** Soient  $A$  et  $B$  deux intervalles compacts de  $\mathbb{R}$ , et  $g : A \rightarrow \mathbb{R}$  une application continue.

(1) Si  $g(A)$  contient  $A$ , alors  $g$  admet un point fixe dans  $A$ .



Olexandr Sharkovsky  
135. 1936-



René Baire  
1874-1932

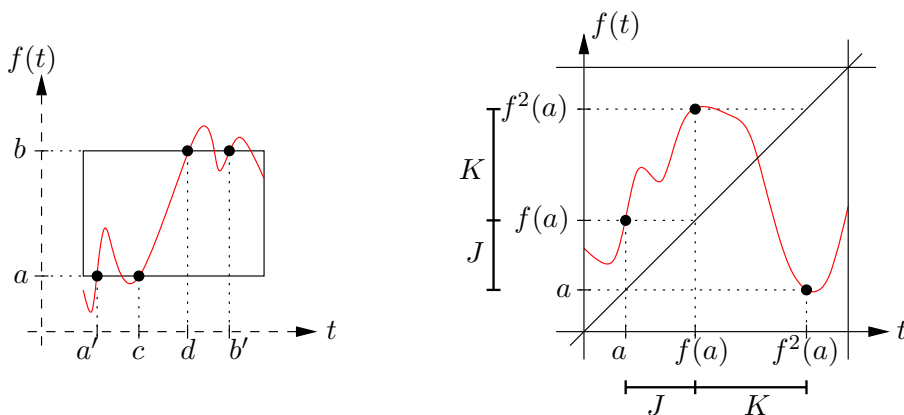


Max Zorn  
1906-1993

(2) Si  $g(A)$  contient  $B$ , alors il existe un sous-intervalle  $C$  de  $A$  tel que  $g(C) = B$ .

**Démonstration.** (1) Le résultat est immédiat si  $A$  est un singleton. Sinon  $A = [a, b]$  avec  $a < b$ . Soient  $a', b' \in A$  tels que  $g(a') = a$ ,  $g(b') = b$ , et supposons par exemple que  $a' < b'$ . Nous avons  $g(a') - a' \leq 0$  et  $g(b') - b' \geq 0$ , donc par le théorème des valeurs intermédiaires, il existe  $x \in [a', b']$  (donc  $x \in A$ ) tel que  $g(x) - x = 0$ .

(2) Le résultat est immédiat si  $B$  est un singleton. Sinon  $B = [a, b]$  avec  $a < b$ . Soient  $a', b' \in A$  tels que  $g(a') = a$ ,  $g(b') = b$ , et supposons par exemple que  $a' < b'$  (voir la figure de gauche ci-dessous). Soient alors  $d = \min\{t \in [a', b'] : g(t) = b\}$ , et  $c = \max\{t \in [a', d] : g(t) = a\}$ , qui existent par compacité et continuité (et le fait que les ensembles considérés sont non vides). Alors par le théorème des valeurs intermédiaires, nous avons comme voulu  $g([c, d]) = [a, b]$ .  $\square$



Pour démontrer le théorème 7.1, fixons  $k \in \mathbb{N} - \{0\}$ . Soit  $a$  un point périodique de période 3. Montrons que si  $a < f(a) < f^2(a)$ , alors  $f$  a un point périodique de période  $k$ .

Quitte à conjuguer  $f$  par la symétrie  $x \mapsto -x$ ,<sup>136</sup> ceci montrera aussi le résultat si nous avons  $a > f(a) > f^2(a)$ , et quitte à remplacer  $a$  par  $f(a)$  ou par  $f^2(a)$ , ceci montrera le théorème.

Posons  $J = [a, f(a)]$  et  $K = [f(a), f^2(a)]$ , de sorte que par continuité  $K \subset f(J)$  et  $(J \cup K) \subset f(K)$  (voir la figure de droite ci-dessus). Par le lemme 7.2 (1) appliqué avec  $A = K$  et  $g = f$ , l'application  $f$  admet un point fixe dans  $K$  (l'intersection avec la diagonale ci-dessus). Nous pouvons donc supposer que  $k \geq 2$ . Posons  $I_n = K$  pour  $0 \leq n < k - 1$ ,  $I_{k-1} = J$  et  $I_k = K$ , de sorte que  $I_{n+1} \subset f(I_n)$  pour  $0 \leq n \leq k - 1$ . Posons  $L_0 = I_0$ , de sorte que  $f^0(L_0) = I_0$  et  $f^1(L_0)$  contient  $I_1$ . Par le lemme 7.2 (2) appliqué avec  $A = L_0$ ,  $B = I_1$  et  $g = f$ , il existe un sous-intervalle  $L_1$  contenu dans  $L_0$  tel que  $f^1(L_1) = I_1$ . Puisque  $k \geq 2$ , ceci implique que  $f^2(L_1)$  contient  $I_2$ . Par récurrence sur  $n$  entre 0 et  $k - 1$ , et par le lemme 7.2 (2) appliqué avec  $A = L_n$ ,  $B = I_{n+1}$  et  $g = f^{n+1}$ , il existe un sous-intervalle  $L_{n+1}$  de  $L_n$  tel que  $f^{n+1}(L_{n+1}) = I_{n+1}$ . En effet, ceci implique que  $f^{n+2}(L_{n+1})$  contient  $I_{n+2}$  tant que  $n \leq k - 2$ , ce qui permet de pousser la récurrence d'un cran tant que  $n \leq k - 2$ .

Par le lemme 7.2 (1), puisque  $f^k(L_k) = I_k = K = L_0$  contient  $L_k$ , l'application  $f^k$  admet un point fixe  $x_k$  dans  $L_k$ . Donc  $x_k$  est un point périodique de  $f$ , dont la période

136. c'est-à-dire quitte à remplacer  $f$  par  $g : x \mapsto -f(-x)$

$p \geq 1$  divise  $k$ . Nous avons

$$f^{k-1}(x_k) \in f^{k-1}(L_k) \subset f^{k-1}(L_{k-1}) = I_{k-1} = J.$$

Supposons par l'absurde que  $p < k$ . Alors nous avons

$$f^{p-1}(x_k) \in f^{p-1}(L_k) \subset f^{p-1}(L_{p-1}) = I_{p-1} = K$$

car  $p-1 < k-1$ . Donc  $f^{k-1}(x_k) = f^{p-1}(f^{k-p}(x_k)) = f^{p-1}(x_k)$  appartient à l'intersection  $J \cap K = \{f(a)\}$ . Par conséquent  $x_k = f(f^{k-1}(x_k)) = f^2(a)$ , donc  $f(x_k) = a$  n'appartient pas à  $K$ . Or nous avons  $f(x_k) \in f(L_k) \subset f(L_1) = I_1$ , qui vaut  $K$  si  $k \neq 2$ . Donc  $k = 2$  et  $p = 1$ , c'est-à-dire que  $x_k$  est un point fixe de  $f$ . Mézalors  $x_k = f(x_k) = a$  est fixe, ce qui contredit le fait que la période de  $a$  soit égale à 3.  $\square$

## 7.2 Récurrence topologique

Soient  $(X, \phi)$  un système dynamique topologique (ou différentiable), et  $x \in X$ .

L'ensemble  $\omega$ -limite de  $x$  est la partie de  $X$ , notée  $\omega(x)$  ou  $\omega_\phi(x)$ , définie par

$$\omega(x) = \bigcap_{T \geq 0} \overline{\{\phi^t(x) : t \geq T\}} = \bigcap_{t \geq 0} \overline{\mathcal{O}^+(\phi^t(x))}.$$

Si  $(X, \phi)$  est à temps discret et inversible, ou est à temps continu, alors nous définissons de même l'ensemble  $\alpha$ -limite de  $x$  pour  $(X, \phi)$ , noté  $\alpha(x)$  ou  $\alpha_\phi(x)$ , comme étant l'ensemble  $\omega$ -limite de  $x$  pour la transformation inverse ou le flot inverse.

Le point  $x$  est dit *positivement récurrent*<sup>137</sup> pour  $(X, \phi)$  si

$$x \in \omega(x).$$

Si  $(X, \phi)$  est à temps discret et inversible, ou est à temps continu, on définit de même un point  $x$  comme étant *négativement récurrent* si  $x \in \alpha(x)$ .

Par exemple, un point périodique est positivement récurrent (et négativement récurrent si  $(X, \phi)$  est à temps discret et inversible, ou est à temps continu), et si  $X$  est séparé (de sorte que toute partie finie de  $X$  soit fermée), alors son ensemble  $\omega$ -limite (et son ensemble  $\alpha$ -limite si  $(X, \phi)$  est à temps discret et inversible, ou est à temps continu), est toute son orbite positive :

$$\omega(x) = \mathcal{O}^+(x).$$

**Remarques.** (1) Si  $X$  est métrisable, alors  $\omega(x)$  est l'ensemble des  $y \in X$  tels qu'il existe une suite de temps  $(t_i)_{i \in \mathbb{N}}$  positive qui converge vers  $+\infty$  telle que  $\lim_{i \rightarrow +\infty} \phi^{t_i}(x) = y$ . Donc si  $X$  est métrisable, alors  $x$  est positivement récurrent si et seulement s'il existe une suite de temps  $(t_i)_{i \in \mathbb{N}}$  positive qui converge vers  $+\infty$  telle que

$$\lim_{i \rightarrow +\infty} \phi^{t_i}(x) = x.$$

(2) Comme intersection décroissante de fermés positivement invariants, tout ensemble  $\omega$ -limite est un fermé positivement invariant, qui est non vide si  $X$  est compact.

137. On prendra bien garde à ne pas confondre un point positivement récurrent (tout court) d'un système dynamique topologique ou différentiable avec un point (infiniment positivement) récurrent **dans une partie mesurable donnée** d'un système dynamique mesurable ou mesuré  $(X, \phi)$ , voir la partie 1.1

**Exercice E.42.** Soit  $(X, \phi)$  un système dynamique topologique (ou différentiable), à temps discret et inversible, ou à temps continu. Montrer que tout ensemble  $\omega$ -limite est globalement invariant.

**Exercice E.43.** Soit  $(X', \phi')$  un système dynamique topologique (ou différentiable), et soit  $h : X \rightarrow X'$  une semi-conjugaison entre  $(X, \phi)$  et  $(X', \phi')$ . Montrer que

$$h(\omega_\phi(x)) \subset \omega_{\phi'}(h(x)), \quad (59)$$

avec égalité si  $h$  est une conjugaison.

### 7.3 L'ensemble (topologiquement) non errant

Soit  $(X, \phi)$  un système dynamique topologique (ou différentiable).

Un point  $x \in X$  est dit (topologiquement) *errant* pour  $(X, \phi)$  s'il existe un voisinage ouvert  $U$  de  $x$  tel que  $\phi^{-t}(U) \cap U$  soit vide pour tout temps  $t \geq 1$ .

L'ensemble des points non errants est appelé *l'ensemble non errant* de  $(X, \phi)$ . Il est noté  $\Omega(\phi)$  (ou  $\Omega(X, \phi)$  lorsqu'il convient de préciser).<sup>138</sup> Le système dynamique topologique  $(X, \phi)$  est dit *non errant* si  $\Omega(X) = X$ .

Comme  $\phi^{-t}(U) \cap U$  est vide si et seulement si  $U \cap \phi^t(U)$  l'est, si  $(X, \phi)$  est à temps discret et inversible, ou s'il est à temps continu, alors  $x$  est errant pour  $(X, \phi)$  si et seulement si l'est pour le système dynamique inverse.

**Proposition 7.3.** *L'ensemble non errant  $\Omega(\phi)$  de  $(X, \phi)$  est un fermé positivement invariant, qui est globalement invariant si  $(X, \phi)$  est à temps discret et inversible, ou à temps continu.*

*Le support de toute mesure (positive borélienne) finie sur  $X$  invariante par  $\phi$  est contenu dans l'ensemble non errant  $\Omega(\phi)$ .*

**Démonstration.** Soit  $\Omega = \Omega(\phi)$ . Montrons que son complémentaire  ${}^c\Omega$  est ouvert. Si  $x \notin \Omega$ , alors il existe un voisinage ouvert  $U$  de  $x$  tel que  $\phi^{-t}(U) \cap U = \emptyset$  pour tout temps  $t \geq 1$ . Donc  $U \subset {}^c\Omega$ , et par conséquent  ${}^c\Omega$  contient un voisinage de chacun de ses points.

Soient  $x \in \Omega$ ,  $t$  un temps positif (ou quelconque si  $(X, \phi)$  est à temps discret et inversible, ou à temps continu) et  $U$  un voisinage de  $\phi^t(x)$ . Alors  $\phi^{-t}(U)$  est un voisinage de  $x$ , donc il existe  $s \geq 1$  tel que  $\phi^{-s-t}(U) \cap \phi^{-t}(U)$  soit non vide. Donc  $\phi^{-s}(U) \cap U$ , qui contient  $\phi^t(\phi^{-s-t}(U) \cap \phi^{-t}(U))$ , est non vide et  $\phi^t(x) \in \Omega$ .

Soit  $\mu$  une mesure sur  $X$  invariante par  $\phi$ . Si  $x \notin \Omega$ , alors il existe un voisinage ouvert  $U$  de  $x$  tel que les ouverts  $\phi^{-n}(U)$  pour  $n \in \mathbb{N}$  soient deux à deux disjoints. Donc

$$\mu(X) \geq \mu\left(\bigcup_{n \in \mathbb{N}} \phi^{-n}(U)\right) = \sum_{n \in \mathbb{N}} \mu(\phi^{-n}(U)) = \sum_{n \in \mathbb{N}} \mu(U).$$

Donc si  $\mu$  est finie, nous avons  $\mu(U) = 0$  et  $x$  n'appartient pas au support de  $\mu$ .  $\square$

**Exemples.** (1) Un point périodique ou positivement récurrent est non errant.

<sup>138.</sup> On prendra bien garde à ne pas confondre l'ensemble non errant (avec un article défini) d'un système dynamique topologique (ou différentiable) avec **un** ensemble non errant (avec un article indéfini) d'un système dynamique mesuré à temps discret  $(X, \mathcal{B}, \mu, \phi)$ , où une partie mesurable  $E \in \mathcal{B}$  est dite *errante* si  $\mu(\phi^{-n}(E) \cap E) = 0$  pour tout  $n \in \mathbb{N} - \{0\}$ .

(2) Puisque l'ensemble non errant  $\Omega(\phi)$  est fermé et contient  $\text{Per}(\phi)$ , il contient l'adhérence  $\overline{\text{Per}(\phi)}$ . En particulier, un système dynamique topologique, dont l'ensemble des points périodiques est dense, comme un système de Bernoulli, est non errant.

(3) Si  $X$  est métrisable, pour tout  $x \in X$ , l'ensemble non errant  $\Omega(\phi)$  contient l'ensemble  $\omega$ -limite de  $x$ . En effet, soient  $y \in \omega(x)$  et  $U$  un voisinage de  $y$  dans  $X$ . Soit  $(t_k)_{k \in \mathbb{N}}$  une suite de temps qui converge vers  $+\infty$  telle que  $\lim_{i \rightarrow +\infty} \phi^{t_i}(x) = y$ . Alors il existe  $i, j \in \mathbb{N}$  tels que  $\phi^{t_i}(x) \in U$ ,  $\phi^{t_j}(x) \in U$  et  $t_j \geq t_i + 1$ . Donc  $x \in \phi^{-t_i}(U) \cap \phi^{-t_j}(U) = \phi^{-t_i}(U \cap \phi^{-(t_j - t_i)}(U))$ , et par conséquent  $U \cap \phi^{-(t_j - t_i)}(U)$  est non vide. L'ensemble non errant  $\Omega(\phi)$  contient donc l'adhérence de la réunion des ensembles  $\omega$ -limites des points de  $X$ .

**Exercice E.44.** Soit  $(X', \phi')$  un système dynamique topologique (ou différentiable), et soit  $h : X \rightarrow X'$  une semi-conjugaison entre  $(X, \phi)$  et  $(X', \phi')$ . Montrer que

$$h^{-1}(\Omega(\phi')) \supset \Omega(\phi), \quad (60)$$

avec égalité si  $h$  est une conjugaison.

## 7.4 Transitivité (topologique), mélange topologique et minimalité (topologique)

Soit  $(X, \phi)$  un système dynamique topologique (ou différentiable).

Ce système dynamique est dit *positivement transitif* si pour tous les ouverts non vides  $U$  et  $V$ , il existe un temps  $t \geq 0$  tel que l'intersection  $U \cap \phi^{-t}(V)$  soit non vide.

Si  $(X, \phi)$  est à temps discret et inversible (respectivement à temps continu), ce système dynamique est dit *transitif*, ou *topologiquement transitif* lorsqu'il convient d'insister sur la nature topologique de la terminologie, si pour tous les ouverts non vides  $U$  et  $V$ , il existe  $t \in \mathbb{Z}$  (respectivement  $t \in \mathbb{R}$ ) tel que l'intersection  $U \cap \phi^{-t}(V)$  soit non vide.

En pratique, nous utiliserons les caractérisations (2) ou (ii) de l'exercice E.46 pour montrer ou utiliser qu'un système dynamique est positivement transitif ou transitif.

**Exercice E.45.** Soit  $h : X \rightarrow X'$  une semi-conjugaison entre  $(X, \phi)$  et un système dynamique topologique  $(X', \phi')$ . Montrer que si  $(X, \phi)$  est positivement transitif (respectivement à temps discret et inversible, ou à temps continu, et transitif), alors le facteur  $(X', \phi')$  l'est aussi.

Remarquons que le système dynamique inversible  $x \mapsto x + 1$  sur l'espace topologique discret  $\mathbb{Z}$  est transitif, mais pas positivement transitif. La réciproque de la propriété évidente que topologiquement transitif implique transitif est néanmoins souvent vraie, comme le montre l'exercice E.49.

**Exercice E.46.** Soient  $X$  un espace de Baire<sup>139</sup> admettant une base dénombrable d'ouverts, et  $\phi : X \rightarrow X$  une application continue. Montrer que les affirmations suivantes sont équivalentes :

- (1) la transformation  $\phi$  est positivement transitive,
- (2) il existe un point dans  $X$  dont l'ensemble  $\omega$ -limite est égal à  $X$ ,

139. Rappelons qu'un *espace de Baire*<sup>135</sup>  $X$  est un espace topologique dans lequel l'intersection de toute famille dénombrable d'ouverts denses est dense. Un *théorème de Baire* dit qu'un espace métrique complet ou un espace localement compact est un espace de Baire. On appelle  $G_\delta$  (dense) dans  $X$  toute intersection dénombrable d'ouverts denses.



(3) l'ensemble  $Y$  des points dans  $X$  dont l'ensemble  $\omega$ -limite est égal à  $X$  est un  $G_\delta$  dense de  $X$ .

Si  $\phi$  est un homéomorphisme, montrer que les affirmations suivantes sont équivalentes :

- (i) la transformation  $\phi$  est transitive,
- (ii) il existe un point dans  $X$  dont l'orbite est dense dans  $X$ ,
- (iii) l'ensemble  $Z$  des points dans  $X$ , dont l'orbite est dense dans  $X$ , est un  $G_\delta$  dense de  $X$ .

Le système dynamique  $(X, \phi)$  est dit *topologiquement mélangeant* si pour tous les ouverts non vides  $U$  et  $V$ , il existe  $T \geq 0$  tel que pour tout  $t \geq T$ , l'intersection  $U \cap \phi^{-t}(V)$  soit non vide. Notons qu'il suffit de vérifier cette condition pour  $U$  et  $V$  dans une base d'ouverts de  $X$ . Il est immédiat qu'un système dynamique topologiquement mélangeant est positivement transitif.

Par un argument similaire à celui de l'exercice E.45 (voir l'exercice E.48), si  $h : X \rightarrow X'$  est une semi-conjugaison entre  $(X, \phi)$  et un système dynamique topologique  $(X', \phi')$ , et si  $(X, \phi)$  est topologiquement mélangeant, alors le facteur  $(X', \phi')$  l'est.

**Exemples.** (1) Une rotation du cercle n'est pas topologiquement mélangeante. En effet, si  $U = \{e^{i\theta} : \theta \in ]0, \frac{\pi}{2}[ \}$  et  $V = \{e^{i\theta} : \theta \in ]\pi, \frac{3\pi}{2}[ \}$ , alors, que la rotation soit rationnelle ou pas, pour tout  $\epsilon > 0$ , il existe des puissances arbitrairement grandes qui envoient  $V$  dans  $\{e^{i\theta} : \theta \in [\pi - \epsilon, \frac{3\pi}{2} + \epsilon] \}$ , donc de manière disjointe à  $U$ .

Une rotation rationnelle du cercle n'est pas transitive, car elle n'admet pas d'orbite dense (toute orbite est périodique). Une rotation irrationnelle du cercle est transitive, car toute orbite est dense.

(2) Un décalage de Bernoulli sur un alphabet dénombrable est positivement transitif (donc transitif dans le cas bilatère).

En effet, soient  $\mathcal{A}$  un alphabet dénombrable et  $(\Sigma_+, \sigma_+)$  le système de Bernoulli unilatère associé. Puisque l'ensemble des mots finis sur  $\mathcal{A}$  est dénombrable, il existe un mot  $x = (x_i)_{i \in \mathbb{N}} \in \Sigma_+$  qui contient tout mot fini. L'orbite positive de  $w$  par le décalage  $\sigma_+$  est dense, puisque pour tout mot  $y \in \Sigma_+$  et pour tout  $N \in \mathbb{N}$ , il existe un entier  $n$  tel que les mots initiaux de longueur  $N$  de  $y$  et  $\sigma_+^n(x)$  coïncident. Donc l'ensemble  $\omega$ -limite de  $x$  est égal à tout  $\Sigma_+$ . Le cas bilatère  $(\Sigma, \sigma)$  se traite de manière similaire en complétant  $x = (x_i)_{i \in \mathbb{N}}$  par un passé arbitraire en  $x' = (x_i)_{i \in \mathbb{Z}} \in \Sigma$ , et en montrant que l'orbite positive de  $x'$  par le décalage  $\sigma$  est dense.

Un décalage de Bernoulli est topologiquement mélangeant (ce qui rédémontre qu'il est positivement transitif).

En effet, (le cas bilatère se traite de manière similaire) pour tous les  $p, p' \in \mathbb{N}$ , et  $x_0, \dots, x_p, y_0, \dots, y_{p'}$  dans l'alphabet  $\mathcal{A}$ , nous avons

$$\begin{aligned} & [x_0, \dots, x_p] \cap \sigma^{-n}([y_0, \dots, y_{p'}]) \\ &= \{(z_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}} : z_0 = x_0, \dots, z_p = x_p, z_n = y_0, \dots, z_{n+p'} = y_{p'}\} \end{aligned}$$

est non vide pour tous les  $n \in \mathbb{N}$  tels que  $n > p$ . Le résultat découle alors du fait que l'ensemble des cylindres initiaux est une base d'ouverts de la topologie de  $\Sigma_+$ .

Le système dynamique  $(X, \phi)$  est dit *positivement minimal* si l'une des conditions équivalentes <sup>140</sup> suivantes est vérifiée :

- (1) toute orbite positive est dense,
- (2) l'ensemble  $\omega$ -limite de tout point de  $X$  est égal à  $X$ ,
- (3) les seuls fermés positivement invariants de  $X$  sont  $\emptyset$  et  $X$ .

Si  $(X, \phi)$  est à temps discret et inversible, ou s'il est à temps continu, il est dit *minimal* si l'une des conditions équivalentes suivantes est vérifiée :

- (1) toute orbite est dense,
- (2) les seuls fermés globalement invariants de  $X$  sont  $\emptyset$  et  $X$ .

Une partie  $Y$  de  $X$  est dite (positivement) *minimale* dans  $(X, \phi)$  si elle est fermée, non vide, (positivement) invariante et minimale pour l'inclusion. Cette dernière condition signifie que  $Y$  ne contient pas de fermé non vide (positivement) invariant autre que  $Y$ , ou de manière équivalente, que la restriction du système dynamique à  $Y$  est (positivement) minimale.

**Exemples.** (i) Un point fixe ou une orbite positivement périodique (respectivement périodique) est une partie positivement minimale (respectivement minimale), si  $X$  est séparé (de sorte que toute partie finie de  $X$  soit fermée)

(ii) Une rotation irrationnelle du cercle est minimale, car toute orbite est dense.

**Remarques.** (1) Supposons  $(X, \phi)$  à temps discret et inversible, ou à temps continu. Si le système dynamique topologique  $(X, \phi)$  est positivement minimal, il est aussi minimal (car les parties globalement invariantes sont aussi positivement invariantes).

Par exemple, le système dynamique topologique inversible  $x \mapsto x+1$  sur l'espace discret  $\mathbb{Z}$  est minimal, mais pas positivement minimal. Mais si  $X$  est compact, alors les propriétés d'être minimal et positivement minimal sont équivalentes. En effet, tout ensemble  $\omega$ -limite est fermé, non vide par compacité et globalement invariant par l'exercice E.42, donc égal à  $X$  si  $(X, \phi)$  est minimal.

(2) Comme la frontière <sup>141</sup> d'une partie (positivement) minimale  $M$  est encore un fermé (positivement) invariant contenu dans  $M$ , elle est

- ou bien vide, auquel cas  $M = X$  si  $X$  est connexe, donc ne s'écrit pas comme réunion de deux fermés non vides disjoints  $\overline{M} = M$  et  ${}^c\overline{M}$ ,
- ou bien égale à  $M$ , qui est donc d'intérieur vide puisque  $\partial M = M - \overset{\circ}{M}$ .

<sup>140</sup>. Soit  $Y$  un fermé positivement invariant, alors pour tout  $y \in Y$ , les fermés positivement invariants  $\overline{\mathcal{O}^+(y)}$  et  $\omega(y)$  sont contenus dans  $Y$ . Donc les deux premières assertions impliquent la troisième. Réciproquement, comme  $\overline{\mathcal{O}^+(x)}$  est un fermé positivement invariant non vide pour tout  $x \in X$ , la troisième assertion implique la première. Enfin, la première assertion implique la deuxième, car alors, pour tout  $x \in X$ , nous avons

$$\omega(x) = \bigcap_{t \geq 0} \overline{\mathcal{O}^+(\phi^t(x))} = \bigcap_{t \geq 0} X = X.$$

<sup>141</sup>. Rappelons que la *frontière* d'une partie  $A$  d'un espace topologique  $X$  est

$$\partial A = \overline{A} \cap \overline{{}^c A} = \overline{A} - \overset{\circ}{A},$$

c'est-à-dire  $\partial A$  est l'ensemble des points de  $X$  dont tout voisinage rencontre à la fois  $A$  et son complémentaire.

(3) Tout système dynamique topologique  $(X, \phi)$ , où  $X$  est séparé, qui contient une orbite périodique et n'est pas réduit à une orbite périodique, n'est pas positivement minimal. Par exemple, un décalage de Bernoulli unilatère sur un alphabet à au moins deux éléments n'est pas positivement minimal, et un décalage de Bernoulli bilatère sur un alphabet à au moins deux éléments n'est pas minimal. Nous verrons ultérieurement que l'espace des phases d'un décalage de Bernoulli sur un alphabet à au moins deux éléments admet de très nombreuses parties minimales (outre l'infinité dénombrable des orbites périodiques, voir l'exemple (2) de la partie 0.1).

**Proposition 7.4.** *Si  $X$  est compact non vide, alors  $X$  contient au moins un minimal.*

**Démonstration.** Soit  $\mathcal{F}$  l'ensemble (non vide car  $X$  est fermé et non vide) des fermés non vides de  $X$  positivement invariants, partiellement ordonné par l'inclusion. Nous voulons montrer que  $\mathcal{F}$  contient un élément minimal, qui sera donc positivement minimal, et minimal par compacité, voir la remarque (1) ci-dessus. Ceci résulte du lemme de Zorn <sup>135</sup> (voir [Kri]), car toute partie non vide  $\mathcal{F}'$  de  $\mathcal{F}$  totalement ordonnée admet un minorant : l'intersection des éléments de  $\mathcal{F}'$ , qui est non vide par compacité.  $\square$

Comme tout point d'une partie positivement minimale est positivement récurrent, le résultat suivant est une conséquence immédiate de l'existence de minimaux.

**Théorème 7.5. (Théorème de récurrence de Birkhoff)** *Soit  $X$  un espace topologique compact. Toute application continue  $\phi : X \rightarrow X$  admet un point positivement récurrent.*  $\square$

**Remarques.** (1) Il peut n'y avoir qu'un seul point positivement récurrent. Par exemple, prendre  $X$  le compactifié d'Alexandrov  $\widehat{\mathbb{R}} \simeq \mathbb{S}_1$  et  $\phi(x) = x + 1$  pour tout  $x \in \mathbb{R}$ , ainsi que  $\phi(\infty) = \infty$ , qui est l'unique point positivement récurrent.

(2) Il peut n'y avoir aucun point périodique. Par exemple, prendre une rotation irrationnelle du cercle.

(3) Donnons une démonstration du théorème 7.5 qui n'utilise pas l'axiome du choix, sous l'hypothèse supplémentaire que  $X$  est métrisable. Elle utilise le théorème de récurrence de Poincaré 1.1.

**Seconde démonstration pour  $X$  métrisable du théorème 7.5.** Soient  $X$  un espace compact métrisable, de tribu borélienne  $\mathcal{B}$ , et  $(U_n)_{n \in \mathbb{N}}$  une base dénombrable d'ouverts de  $X$ . Soit  $\phi : X \rightarrow X$  une application continue. Par la proposition 0.6, il existe une mesure de probabilité  $\mu$  invariante par  $\phi$ . Notons  $R_+$  l'ensemble mesurable des  $x \in X$  tels que, pour tout  $m \in \mathbb{N}$ , si  $x \in U_m$ , alors l'orbite positive  $(\phi^n(x))_{n \in \mathbb{N}}$  repasse une infinité de fois dans  $U_m$ . Puisque  $X = \bigcup_{n \in \mathbb{N}} U_n$ , le théorème de récurrence de Poincaré 1.1 appliqué à  $(X, \mathcal{B}, \mu)$  et un argument de dénombrabilité montre que  $R_+$  est de mesure totale dans  $X$ , et en particulier non vide. Mais si  $x \in R_+$ , pour tout voisinage ouvert  $U$  de  $x$ , il existe  $m \in \mathbb{N}$  tel que  $x \in U_m \subset U$ , donc  $x$  revient une infinité de fois dans  $U$ . D'où  $x$  est positivement récurrent, ce qui démontre le théorème de récurrence de Birkhoff.  $\square$

**Remarque.** Cette démonstration montre que si  $X$  est métrisable compact, alors l'ensemble des points positivement récurrents est de mesure totale pour toute mesure de probabilité  $\mu$  invariante.

Supposons de plus  $\phi$  inversible. Notons  $R_-$  l'ensemble mesurable des  $x \in X$  tels que pour tout  $m \in \mathbb{N}$ , si  $x \in U_m$ , alors l'orbite négative  $(\phi^{-n}(x))_{n \in \mathbb{N}}$  repasse une infinité de fois dans  $U_m$ . En considérant  $R_- \cap R_+$ , nous obtenons alors que l'ensemble des points récurrents de  $(X, \phi)$  est de mesure totale pour toute mesure de probabilité  $\mu$  invariante.

**Exercice E.47.** Soient  $p \in \mathbb{Z} - \{-1, 0, 1\}$ ,  $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$  et  $\phi_p : \mathbb{T}^1 \rightarrow \mathbb{T}^1$  l'application <sup>142</sup> continue définie par  $x \mapsto px$ .

(1) Montrer que le système dynamique topologique à temps discret  $(\mathbb{T}^1, \phi_p)$  est non inversible, qu'il est conjugué au système dynamique topologique  $(\mathbb{S}_1, z \mapsto z^p)$ , que l'ensemble  $\text{Per}(\phi_p)$  des points périodiques de  $\phi_p$  est dense dans le cercle  $\mathbb{T}^1$ , que le système dynamique topologique  $(\mathbb{T}^1, \phi_p)$  est non errant, qu'il est positivement transitif, et qu'il est topologiquement mélangeant.

(2) Supposons  $p \geq 2$ . Soit  $\Theta : \{0, \dots, p-1\}^{\mathbb{N}} \rightarrow \mathbb{T}^1$  l'application définie par

$$(x_i)_{i \in \mathbb{N}} \mapsto \sum_{i \in \mathbb{N}} \frac{x_i}{p^{i+1}} \pmod{\mathbb{Z}}.$$

Montrer que  $\Theta$  est une semi-conjugaison entre le système de Bernoulli unilatère d'alphabet  $\{0, \dots, p-1\}$  et le système dynamique topologique  $(\mathbb{T}^1, \phi_p)$ , que l'on appelle un codage de  $(\mathbb{T}^1, \phi_p)$ . Retrouver les propriétés précédentes.

(3) Supposons toujours  $p \geq 2$ . Notons  $C^0(\mathbb{T}^1)$  l'espace de Banach des applications continues réelles sur  $\mathbb{T}^1$ , muni de la norme uniforme  $\|\cdot\|_\infty$ . Notons  $L : C^0(\mathbb{T}^1) \rightarrow C^0(\mathbb{T}^1)$  l'application linéaire continue (de norme au plus 1), appelée opérateur de transfert, définie, pour tous les  $f \in C^0(\mathbb{T}^1)$  et  $y \in \mathbb{T}^1$ , par

$$Lf : y \mapsto \frac{1}{\text{Card}(\phi_p^{-1}(y))} \sum_{z \in \phi_p^{-1}(y)} f(z) = \frac{1}{p} \sum_{i=0}^{p-1} f\left(\frac{y+i}{p}\right),$$

en notant par abus de la même manière  $y$  et un relevé de  $y$  dans  $\mathbb{R}$ , le terme de droite ne dépendant pas du choix de ce relevé. Montrer qu'en notant  $\lambda$  la mesure de Lebesgue <sup>143</sup> de  $\mathbb{T}^1$ , l'application  $\phi_p$  préserve la mesure de Lebesgue  $\lambda$ , et que nous avons, pour tous les  $f, g \in C^0(\mathbb{T}^1)$ ,

$$\int_{\mathbb{T}^1} (f \circ \phi_p) g \, d\lambda = \int_{\mathbb{T}^1} f (Lg) \, d\lambda. \quad (61)$$

Montrer que pour tous les  $f \in C^1(\mathbb{T}^1)$  et  $n \in \mathbb{N}$ , nous avons

$$\left| L^n f(x) - \int_0^1 f(t) dt \right| \leq \frac{1}{p^n} \int_0^1 |f'(t)| dt. \quad (62)$$

En déduire que la transformation  $\phi_p$  est mélangeante pour la mesure de Lebesgue du cercle, c'est-à-dire (voir aussi la partie 4.5) que pour tous les  $f, g \in L^2(\mathbb{T}^1, \lambda)$ , nous avons

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{T}^1} (f \circ \phi_p^n) g \, d\lambda = \left( \int_{\mathbb{T}^1} f \, d\lambda \right) \left( \int_{\mathbb{T}^1} g \, d\lambda \right).$$

En déduire que la transformation  $\phi_p$  est exponentiellement mélangeante pour la mesure de Lebesgue du cercle (voir la partie 4.5), et plus précisément que pour tous les  $f, g \in C^1(\mathbb{T}^1)$ ,

142. appelée l'application de doublement de l'angle si  $p = 2$

143. Par définition (voir aussi le début du chapitre 4), pour toute fonction continue  $f \in C^0(\mathbb{T}^1)$  sur  $\mathbb{T}^1$ , identifiée à une fonction 1-périodique sur  $\mathbb{R}$ , pour tout  $t_0 \in \mathbb{R}$ , nous avons

$$\int_{\mathbb{T}^1} f \, d\lambda = \int_{t_0}^{t_0+1} f(t) \, dt.$$

en définissant la norme de Sobolev  $W^{1,\infty}$  de  $f$  par  $\|f\|_{1,\infty} = \max\{\|f\|_\infty, \|f'\|_\infty\}$ , nous avons

$$\left| \int_{\mathbb{T}^1} (f \circ \phi_p^n) g \, d\lambda - \left( \int_{\mathbb{T}^1} f \, d\lambda \right) \left( \int_{\mathbb{T}^1} g \, d\lambda \right) \right| \leq \frac{1}{p^n} \|f\|_{1,\infty} \|g\|_{1,\infty}. \quad (63)$$

## 7.5 Récurrence multiple

Voici une généralisation du théorème de récurrence de Birkhoff 7.5.

**Théorème 7.6. (Furstenberg-Weiss<sup>144</sup>)** Soient  $X$  un espace topologique métrisable compact non vide et  $f : X \rightarrow X$  une application continue. Alors, pour tout  $\ell$  dans  $\mathbb{N} - \{0\}$ , il existe un point  $x$  de  $X$  et une suite strictement croissante d'entiers  $(n_k)_{k \in \mathbb{N}}$  tels que

$$\forall p \in \{1, \dots, \ell\}, \quad \lim_{k \rightarrow \infty} f^{pn_k}(x) = x.$$

Autrement dit, il existe un point  $x$  positivement récurrent pour  $f, f^2, \dots, f^\ell$ , avec les mêmes temps de presque-retour. Nous utiliserons le lemme indépendant suivant.

**Lemme 7.7.** Soient  $Y$  un espace topologique métrisable complet et  $\varphi : Y \rightarrow \mathbb{R}$  une fonction semi-continue supérieurement<sup>145</sup>. Alors l'ensemble des points de continuité de  $\varphi$  est dense dans  $Y$ .

**Démonstration.** Quitte à remplacer  $\varphi$  par  $e^\varphi$ , nous pouvons supposer que  $\varphi$  est minorée. Notons, pour tout  $\epsilon > 0$ ,

$$Y_\epsilon = \{y \in Y : \exists (y_k)_{k \in \mathbb{N}} \in Y^{\mathbb{N}}, \lim_{k \rightarrow \infty} y_k = y \text{ et } \limsup_{k \rightarrow \infty} \varphi(y_k) \leq \varphi(y) - \epsilon\}.$$

Comme  $\varphi$  est semi-continue supérieurement, l'ensemble  $Y_\epsilon$  est fermé<sup>146</sup> et l'ensemble des points de continuité de  $\varphi$  est l'intersection des ouverts  ${}^c Y_{\frac{1}{n}}$  pour  $n \in \mathbb{N}$ . Par le théorème de Baire, il suffit de voir que chacun de ces ouverts est dense, c'est-à-dire que chacun des fermés  $Y_\epsilon$  est d'intérieur vide.



Benjamin Weiss  
1941-



Bartel  
van der Waerden  
1903-1996



Arnaud Denjoy  
1884-1974

144.

145. Une fonction  $\varphi : Y \rightarrow \mathbb{R}$  définie sur un espace topologique  $Y$  est *semi-continue supérieurement* si pour tout  $y_0$  dans  $Y$ , nous avons  $\limsup_{y \rightarrow y_0} \varphi(y) \leq \varphi(y_0)$ , voir par exemple [Bre, page 8], [Pau3, §5.4]. Par exemple, une borne inférieure de fonctions continues est semi-continue supérieurement.

146. En effet, soit  $(y^i)_{i \in \mathbb{N}}$  une suite dans  $Y_\epsilon$  qui converge vers  $y^\infty$  dans  $Y$ . Montrons que  $y^\infty \in Y_\epsilon$ . Pour tout  $i \in \mathbb{N}$ , considérons une suite  $(y_k^i)_{k \in \mathbb{N}} \in Y^{\mathbb{N}}$  qui vérifie  $\lim_{k \rightarrow \infty} y_k^i = y^i$  et  $\limsup_{k \rightarrow \infty} \varphi(y_k^i) \leq \varphi(y^i) - \epsilon$ . Pour tout  $i \in \mathbb{N}$ , soit  $k_i \in \mathbb{N}$  tel que  $k_i \geq i$  et  $\varphi(y_{k_i}^i) \leq \varphi(y^i) - \epsilon + \frac{1}{1+i}$ . Alors  $\lim_{i \rightarrow \infty} y_{k_i}^i = y^\infty$  et, par la semi-continuité supérieure de  $\varphi$ , nous avons

$$\limsup_{i \rightarrow \infty} \varphi(y_{k_i}^i) \leq \limsup_{i \rightarrow \infty} \varphi(y^i) - \epsilon \leq \varphi(y^\infty) - \epsilon.$$

Si, par l'absurde, l'intérieur de  $Y_\epsilon$  contenait un point  $y_0$ , il contiendrait aussi un point  $y_1$  tel que  $\varphi(y_1) \leq \varphi(y_0) - \frac{\epsilon}{2}$ . Nous construirions ainsi une suite de points  $(y_k)_{k \in \mathbb{N}}$  dans l'intérieur de  $Y_\epsilon$  telle que  $\varphi(y_k) \leq \varphi(y_{k-1}) - \frac{\epsilon}{2}$ . Ceci contredirait le fait que  $\varphi$  est minorée.  $\square$

**Démonstration du théorème de Furstenberg-Weiss.** Par la proposition 7.4, nous pouvons supposer que  $X$  est minimal. Procédons par récurrence sur  $\ell$ . Pour  $\ell = 1$ , c'est le théorème de récurrence de Birkhoff 7.5. Supposons le résultat vrai pour  $\ell - 1$ . Notons  $x$  un point de  $X$  et  $(n_k)_{k \in \mathbb{N}}$  une suite strictement croissante d'entiers tels que

$$\forall p \in \{1, \dots, \ell - 1\}, \quad \lim_{k \rightarrow \infty} f^{pn_k}(x) = x.$$

Introduisons l'espace produit  $Y = X^\ell$ , la diagonale  $\Delta = \{(x, \dots, x) : x \in X\}$  de  $Y$  et les deux transformations continues  $g$  et  $h$  de  $Y$  données par

$$g(x_1, \dots, x_\ell) = (f(x_1), f^2(x_2), \dots, f^\ell(x_\ell))$$

$$h(x_1, \dots, x_\ell) = (f(x_1), f(x_2), \dots, f(x_\ell)).$$

Nous voulons trouver un point positivement récurrent de  $g$  qui soit sur  $\Delta$ . Pour cela, nous utiliserons les trois propriétés suivantes de  $g$  et  $h$ .

**Étape 1.**  $\left\{ \begin{array}{l} \text{(i) Les applications } g \text{ et } h \text{ commutent (c'est-à-dire } g \circ h = h \circ g). \\ \text{(ii) La transformation } h \text{ préserve } \Delta \text{ et } \Delta \text{ est un minimal pour } h. \\ \text{(iii) Il existe } y_0 \text{ dans } \Delta \text{ et une suite de points } (z_{0,k})_{k \in \mathbb{N}} \text{ dans } \Delta \text{ tels} \\ \text{que} \\ \lim_{k \rightarrow \infty} g^{n_k}(z_{0,k}) = y_0. \end{array} \right.$

**Démonstration.** Les points (i) et (ii) sont clairs. Pour vérifier le point (iii), nous prenons  $y_0 = (x, \dots, x)$  et  $z_{0,k} = (x_k, \dots, x_k)$  où  $x_k \in X$  vérifie  $f^{n_k}(x_k) = x$ . C'est possible, car par minimalité et compacité,  $f$  est surjective. Nous avons alors

$$g^{n_k}(z_{0,k}) = (x, f^{n_k}(x), \dots, f^{(\ell-1)n_k}(x)) \xrightarrow{k \rightarrow \infty} (x, x, \dots, x) = y_0. \quad \square$$

**Étape 2.**  $\forall y \in \Delta, \forall \epsilon > 0, \exists z \in \Delta, \exists n \geq 1, \quad d(g^n(z), y) < \epsilon.$

**Démonstration.** Soient  $y \in \Delta$  et  $\epsilon > 0$ . Par minimalité de  $\Delta$  pour  $h$ , il existe  $m \geq 0$  tel que  $d(h^m(y_0), y) < \epsilon$ . Nous prenons alors  $z = h^m(z_{0,k})$  et  $n = n_k$  pour  $k$  suffisamment grand, car

$$g^{n_k}(z) = h^m g^{n_k}(z_{0,k}) \xrightarrow{k \rightarrow \infty} h^m(y_0). \quad \square$$

**Étape 3.**  $\forall \epsilon > 0, \exists y \in \Delta, \exists n \geq 1, \quad d(g^n(y), y) < \epsilon.$

**Démonstration.** Soient  $\epsilon > 0$ ,  $\epsilon_0 = \frac{\epsilon}{2}$  et  $y_0$  un point quelconque de  $\Delta$ . L'étape 2 permet de trouver un point  $y_1$  dans  $\Delta$  et  $n_1 \geq 1$  tel que  $d(g^{n_1}(y_1), y_0) < \epsilon_0$ . La continuité de  $g^{n_1}$  permet de trouver  $\epsilon_1 < \epsilon_0$  tel que

$$\forall y \in \Delta, \quad d(y, y_1) < \epsilon_1 \Rightarrow d(g^{n_1}(y), y_0) < \epsilon_0.$$

Recommençons avec  $\epsilon_1$  et  $y_1$ . Nous choisissons ainsi, de proche en proche, pour tout  $k \geq 1$  des entiers  $n_k \geq 1$ , des points  $y_k$  dans  $\Delta$  et des réels  $\epsilon_k < \epsilon_{k-1}$  tels que

$$\forall y \in \Delta, \quad d(y, y_k) < \epsilon_k \Rightarrow d(g^{n_k}(y), y_{k-1}) < \epsilon_{k-1}.$$

Nous avons alors

$$\forall i < j, \quad d(g^{n_j+n_{j-1}+\dots+n_{i+1}}(y_j), y_i) < \epsilon_i < \frac{\epsilon}{2}.$$

Par la compacité de  $\Delta$ , nous pouvons trouver  $i < j$  tels que  $d(y_i, y_j) < \frac{\epsilon}{2}$ . Nous prenons alors  $y = y_j$  et  $n = n_j + n_{j-1} + \dots + n_{i+1}$ . Par l'inégalité triangulaire, nous avons donc  $d(g^n(y), y) < \epsilon$ .  $\square$

Pour conclure la démonstration du théorème de Furstenberg-Weiss, introduisons la fonction  $\varphi : \Delta \rightarrow [0, +\infty[$  définie par

$$\varphi(y) = \inf_{n \in \mathbb{N} - \{0\}} d(y, g^n(y)).$$

Cette fonction  $\varphi$  est semi-continue supérieurement, car pour tout  $n \in \mathbb{N} - \{0\}$ , la fonction  $y \mapsto d(y, g^n(y))$  est continue. D'après le lemme 7.7 ci-dessus, il existe un point de continuité  $a$  de  $\varphi$ . Il nous suffit de montrer que  $\varphi(a) = 0$  pour obtenir le résultat de récurrence au rang  $k$ .

Supposons, par l'absurde, que  $\varphi(a) > 0$ . Il existe alors un réel  $\delta > 0$  et un ouvert non vide  $V$  de  $\Delta$  tel que, pour tout  $y$  dans  $V$ , nous avons  $\varphi(y) > \delta$ . Nous avons l'égalité

$$\bigcup_{m \geq 1} h^{-m}(V) = \Delta$$

car le complémentaire de cette union (non vide) est un fermé positivement  $h$ -invariant, donc vide par minimalité de  $\Delta$  pour  $h$ . Par compacité de  $\Delta$ , nous pouvons trouver une partie finie  $F$  de  $\mathbb{N} - \{0\}$  telle que nous ayons

$$\bigcup_{m \in F} h^{-m}(V) = \Delta.$$

Par l'uniforme continuité des applications  $h^m$  sur  $\Delta$ , il existe  $\epsilon > 0$  tel que

$$\forall y', y'' \in \Delta, \quad d(y', y'') < \epsilon \Rightarrow \forall m \in F, \quad d(h^m(y'), h^m(y'')) < \delta.$$

Considérons alors le point  $y$  de  $\Delta$  et l'entier  $n \geq 1$  donnés par l'étape 3. Nous avons donc  $d(g^n(y), y) < \epsilon$ . Ce point  $y$  est dans un des ouverts  $h^{-m}(V)$  pour un  $m$  dans  $F$ . Nous avons donc  $d(g^n(h^m(y)), h^m(y)) < \delta$ . En particulier, le point  $h^m y$  de  $V$  vérifie  $\varphi(h^m(y)) < \delta$ . Ceci est une contradiction.  $\square$

Voici une application du théorème de Furstenberg-Weiss.

**Théorème 7.8. (van der Waerden<sup>144</sup>)** *Soit  $\mathbb{N} = B_1 \sqcup \dots \sqcup B_q$  une partition finie de  $\mathbb{N}$ , alors il existe une partie  $B_j$  qui contient une suite arithmétique  $\{a, a+n, a+2n, \dots, a+\ell n\}$  avec  $n \geq 1$ , de longueur  $\ell + 1$  aussi grande que l'on veut.*

**Démonstration.** Il s'agit de voir que, pour  $\ell \geq 1$  fixé, une des parties  $B_j$  contient une suite arithmétique de longueur  $\ell + 1$ . La partition ci-dessus définit un mot  $\omega = (\omega_i)_{i \in \mathbb{N}}$  du système de Bernoulli unilatère  $(\Sigma_+ = \{1, \dots, q\}^{\mathbb{N}}, \sigma_+)$  par

$$\forall i \in \mathbb{N}, \quad \omega_i = j \Leftrightarrow i \in B_j.$$

L'espace topologique produit  $\{1, \dots, q\}^{\mathbb{N}}$  est métrisable compact (voir la formule (4)). Soit  $X = \overline{\{\sigma_+^n \omega : n \in \mathbb{N}\}}$  l'adhérence de l'orbite positive de  $\omega$  par le décalage  $\sigma_+$ , qui est positivement invariante par  $\sigma_+$ . D'après le théorème de Furstenberg-Weiss, il existe un point  $\omega'$  de  $X$  et un entier  $n \geq 1$  tel que  $d(\omega', \sigma_+^j \omega') < 1$  pour tout  $j = 1, \dots, \ell$ , c'est-à-dire, tel que  $\omega'_0 = \omega'_n = \omega'_{2n} = \dots = \omega'_{\ell n}$ . Comme  $\omega'$  est dans l'adhérence de l'orbite de  $\omega$ , il existe un entier  $a \geq 0$  tel que  $d(\omega', \sigma_+^a \omega) < e^{-\ell n}$ , c'est-à-dire, tel que  $\omega_{a+i} = \omega'_i$  pour tout  $i = 1, \dots, \ell n$ . En particulier, nous avons  $\omega_a = \omega_{a+n} = \dots = \omega_{a+\ell n}$ . Autrement dit, en notant  $j$  cet entier, la suite arithmétique  $\{a, a+n, a+2n, \dots, a+\ell n\}$  de longueur  $\ell+1$  est contenue dans  $B_j$ .  $\square$

## 7.6 Exercices

**Exercice E.48.** Soient  $(X, \phi)$  et  $(X', \phi')$  deux systèmes dynamiques topologiques, et soit  $h : X \rightarrow X'$  une semi-conjugaison entre  $(X, \phi)$  et  $(X', \phi')$ . Parmi les propriétés suivantes, quelles sont celles qui sont vérifiées par  $(X', \phi')$  si elles le sont par  $(X, \phi)$ ? Donner des contre-exemples lorsque ce n'est pas le cas.

- avoir l'ensemble de ses points périodiques dense,
- avoir un point positivement récurrent,
- ne pas avoir de point errant,
- être positivement transitif (et transitif dans le cas inversible),
- être topologiquement mélangeant,
- être positivement minimal (et minimal dans le cas inversible).

**Exercice E.49.** Soit  $X$  un espace topologique métrisable séparable complet sans point isolé, et soit  $\phi : X \rightarrow X$  un homéomorphisme transitif.

- (1) Montrer que le système dynamique topologique  $(X, \phi)$  est non errant.
- (2) Montrer que pour tout ouvert  $V$  de  $X$ , les ensembles  $W = \bigcup_{n \in \mathbb{N}} \phi^{-n}(V)$  et  $\phi(W)$  ont la même adhérence.
- (3) Montrer que  $\phi$  est positivement transitif.

**Exercice E.50.** Soient  $X$  un espace topologique,  $\phi : X \rightarrow X$  une application continue et  $\mu$  une mesure de probabilité borélienne invariante par  $\phi$  sur  $X$ , de support  $X$ . Montrer que si le système dynamique mesuré  $(X, \mu, \phi)$  est mélangeant (voir la partie 3), alors le système dynamique topologique  $(X, \phi)$  est topologiquement mélangeant.

Montrer qu'une rotation du cercle n'est pas mélangeante pour la mesure de Lebesgue du cercle.

**Exercice E.51.**

- (1) (**Transformation presque périodique**) Soient  $X$  un espace métrique compact non vide, et  $\phi : X \rightarrow X$  un homéomorphisme. Montrer que si le système dynamique topologique  $(X, \phi)$  est topologiquement transitif, et si la suite des itérés  $(\phi^n)_{n \in \mathbb{N}}$  est équicontinue,<sup>147</sup> alors le système dynamique topologique  $(X, \phi)$  est uniquement ergodique.

147. Rappelons (voir par exemple [Pau3, §5.5]) que si  $X$  et  $Y$  sont deux espaces métriques, une famille  $(f_\alpha)_{\alpha \in A}$  d'applications de  $X$  dans  $Y$  est dite *équicontinue* si

$$\forall \epsilon > 0, \forall x \in X, \exists \delta > 0, \forall \alpha \in A, \forall y \in X, \text{ si } d(y, x) \leq \delta \text{ alors } d(f_\alpha(y), f_\alpha(x)) \leq \epsilon.$$

Le *théorème d'Arzela-Ascoli* dit que si  $X, Y$  sont des espaces métriques, si  $(f_\alpha)_{\alpha \in A}$  est une famille d'applications continues de  $X$  dans  $Y$ , telle que



- (2) **(Critère d'unique ergodicité des translations à gauche)** Soit  $G$  un groupe topologique compact et  $\tau$  une translation à gauche sur  $G$ . Montrer que si  $\tau$  est topologiquement transitive, alors le système dynamique topologique  $(G, \tau)$  est uniquement ergodique.
- (3) **(Transformation de l'odomètre  $p$ -adique)** Si  $p$  est un nombre premier, notons  $v_p : \mathbb{Q} \rightarrow \mathbb{Z} \cup \{+\infty\}$  la valuation  $p$ -adique, définie par  $v_p(0) = +\infty$  et si  $x \in \mathbb{Q} \setminus \{0\}$  alors  $v_p(x)$  est l'unique élément  $n \in \mathbb{Z}$  tel que  $x = p^n \frac{r}{s}$  avec  $r, s \in \mathbb{Z} \setminus \{0\}$  non divisibles par  $p$ . Notons  $\mathbb{Z}_p$  l'espace métrique complété de  $\mathbb{Z}$  pour la distance  $p$ -adique  $(x, y) \mapsto |x - y|_p = p^{-v_p(x-y)}$ . Montrer que l'addition de  $\mathbb{Z}$  s'étend en une loi de groupe abélien compact sur  $\mathbb{Z}_p$ , et que la transformation  $\phi : x \mapsto x + 1$  est uniquement ergodique sur  $\mathbb{Z}_p$ .

## 7.7 Indications pour la résolution des exercices

**Correction de l'exercice E.42.** Supposons  $(X, \phi)$  à temps discret et inversible (la démonstration est analogue dans le cas à temps continu). Soient  $x \in X$ ,  $y \in \omega(x)$  et  $n \in \mathbb{Z}$ . Montrons que  $\phi^n(y) \in \omega(x)$ , ce qui donne le résultat.

Puisque  $y \in \omega(x)$ , pour tout  $m \in \mathbb{N}$ , si  $m \geq -n$ , alors  $y \in \overline{\mathcal{O}^+(\phi^{m-n}(x))}$  et puisque  $\phi^n$  est continue et envoie orbite positive sur orbite positive, nous avons <sup>148</sup>

$$\phi^n(y) \in \phi^n(\overline{\mathcal{O}^+(\phi^{m-n}(x))}) \subset \overline{\mathcal{O}^+(\phi^n(\phi^{m-n}(x)))} = \overline{\mathcal{O}^+(\phi^m(x))}.$$

Ceci étant vrai pour tout  $m$  assez grand, nous avons  $\phi^n(y) \in \omega(x)$ , comme souhaité.

**Correction de l'exercice E.43.** Supposons  $(X, \phi)$  à temps discret (la démonstration est analogue dans le cas à temps continu). Nous avons, par la continuité de  $h$  pour la deuxième inclusion,

$$\begin{aligned} h(\omega_\phi(x)) &= h\left(\bigcap_{N \geq 0} \overline{\{\phi^n(x) : n \geq N\}}\right) \subset \bigcap_{N \geq 0} h\left(\overline{\{\phi^n(x) : n \geq N\}}\right) \\ &\subset \bigcap_{N \geq 0} \overline{h(\{\phi^n(x) : n \geq N\})} = \bigcap_{N \geq 0} \overline{\{(\phi')^n(h(x)) : n \geq N\}} \\ &= \omega_{\phi'}(h(x)). \end{aligned}$$

Si  $h$  est une conjugaison, l'inclusion réciproque s'obtient en remplaçant  $h$  par  $h^{-1}$  et  $x$  par  $h(x)$  : nous avons

$$h^{-1}(\omega_{\phi'}(h(x))) \subset \omega_\phi(h^{-1}(h(x))) = \omega_\phi(x),$$

donc  $\omega_{\phi'}(h(x)) \subset h(\omega_\phi(x))$ .

**Correction de l'exercice E.44.** Rappelons que les images réciproques commutent avec les opérations booléennes d'intersection, de réunion et de passage au complémentaire. Soit  $x \in X$  tel que  $h(x) \in {}^c\Omega(\phi')$ . Il s'agit de montrer que  $x \in {}^c\Omega(\phi)$ .

- la famille  $(f_\alpha)_{\alpha \in A}$  est équicontinue,
- pour tout  $x$  dans  $X$ , l'ensemble  $\{f_\alpha(x) : \alpha \in A\}$  est d'adhérence compacte dans  $Y$ ,

alors l'adhérence de  $\{f_\alpha : \alpha \in A\}$  est compacte (et équicontinue) dans l'ensemble  $C(X, Y)$  des applications continues de  $X$  dans  $Y$ , muni de la topologie de la convergence uniforme sur les compacts.

148. Rappelons que si  $g : Y \rightarrow Z$  est une application continue entre deux espaces topologiques, et si  $Y'$  est une partie de  $Y$ , alors  $g(\overline{Y'}) \subset \overline{g(Y')}$ .

Puisque  $h(x)$  est errant, il existe un voisinage  $U$  de  $h(x)$  dans  $X'$  tel que  $(\phi')^{-t}(U) \cap U$  est vide pour tout  $t \geq 1$ . Donc pour tout  $t \geq 1$ , l'intersection

$$\phi^{-t}(h^{-1}(U)) \cap h^{-1}(U) = h^{-1}((\phi')^{-t}(U) \cap U)$$

est vide. Or  $h^{-1}(U)$  est un voisinage de  $x$ , puisque  $h$  est continue. Ceci montre que  $x$  est errant, comme souhaité.

Si  $h$  est une conjugaison, l'inclusion réciproque se démontre comme dans l'exercice qui précède.

**Correction de l'exercice E.45.** Supposons  $(X, \phi)$  positivement transitif (respectivement transitif). Pour tous les ouverts non vides  $U$  et  $V$  de  $X'$ , puisque  $h^{-1}(U)$  et  $h^{-1}(V)$  sont des ouverts non vides de  $X$ , il existe un temps  $t \geq 0$  (respectivement un temps  $t$ ) tel que l'intersection  $h^{-1}(U) \cap \phi^{-t}(h^{-1}(V))$  soit non vide. Comme l'image d'une intersection est contenue dans l'intersection des images, comme

$$\phi^{-t}(h^{-1}(V)) = (h \circ \phi^t)^{-1}(V) = ((\phi')^t \circ h)^{-1}(V) = h^{-1}((\phi')^{-t}(V)),$$

l'intersection  $U \cap (\phi')^{-t}(V)$ , qui contient  $h(h^{-1}(U) \cap h^{-1}((\phi')^{-t}(V)))$ , est non vide. Donc  $(X', \phi')$  est positivement transitif (respectivement transitif).

**Correction de l'exercice E.46.** Considérons la première équivalence triple. Montrons que (2) implique (1). Soit  $x \in X$  tel que  $\omega(x) = X$ . Soient  $U$  et  $V$  deux ouverts non vides. Puisque  $\omega(x) = X$ , il existe  $m \in \mathbb{N}$  tel que  $\phi^m(x) \in U$  et  $n > m$  tel que  $\phi^n(x) \in V$ . Donc  $U \cap \phi^{m-n}(V)$  n'est pas vide, car il contient  $\phi^m(x) = \phi^{m-n}(\phi^n(x))$ .

Il est immédiat que (3) implique (2), donc montrons que (1) implique (3). Soit  $(U_i)_{i \in \mathbb{N}}$  une base dénombrable d'ouverts de  $X$ , et nous pouvons supposer que les  $U_i$  sont non vides. Pour tous les  $i, m \in \mathbb{N}$ , puisque  $\phi$  est positivement transitif, l'ouvert  $\phi^{-m}(U_i)$  est non vide, et  $\bigcup_{n \geq m} \phi^{-n}(U_i)$  est un ouvert dense de  $X$ . Donc l'intersection dénombrable

$$Y = \bigcap_{i, m \in \mathbb{N}} \bigcup_{n \geq m} \phi^{-n}(U_i)$$

est un  $G_\delta$  dense de  $X$ .

Montrons que pour tout  $x \in X$ , nous avons  $\omega(x) = X$  si et seulement si  $x \in Y$ . Le sens direct est immédiat, car si  $\omega(x) = X$ , alors pour tous les  $i, m \in \mathbb{N}$ , il existe  $n \geq m$  tel que  $\phi^n(x) \in U_i$ .

Réciproquement, soit  $x \in Y$ , montrons que  $\omega(x) = X$ . Soient  $x' \in X$  et  $U$  un voisinage de  $x'$ . Puisque  $(U_i)_{i \in \mathbb{N}}$  est une base d'ouverts, il existe  $i \in \mathbb{N}$  tel que  $U_i \subset U$ . Puisque  $x \in Y$ , pour tout  $m \in \mathbb{N}$ , il existe  $n \geq m$  tel que  $\phi^n(x) \in U_i \subset U$ . Donc  $x' \in \omega(x)$  et  $\omega(x) = X$ .

La seconde équivalence triple se démontre de manière analogue, en considérant

$$Z = \bigcap_{i \in \mathbb{N}} \bigcup_{n \in \mathbb{Z}} \phi^{-n}(U_i).$$

**Correction de l'exercice E.47.** (1) L'application  $\phi_p : x \mapsto px$  n'est pas injective, car 0 et  $\frac{1}{p}$  ont même image. L'homéomorphisme  $t \bmod \mathbb{Z} \mapsto e^{2i\pi t}$  de  $\mathbb{R}/\mathbb{Z}$  dans  $\mathbb{S}_1$  est une conjugaison entre  $x \mapsto px$  et  $z \mapsto z^p$ . L'ensemble

$$\left\{ \frac{k}{p^n - 1} : n \in \mathbb{N} - \{0\}, k \in \mathbb{Z} \right\}$$

est dense dans  $\mathbb{R}$ , donc son image dans  $\mathbb{R}/\mathbb{Z}$  par la projection canonique est dense. De plus, cette image est constituée de points périodiques pour  $\phi_p$ , car

$$p^n \frac{k}{p^n - 1} = \frac{(p^n - 1)k}{p^n - 1} + \frac{k}{p^n - 1} = \frac{k}{p^n - 1} \pmod{\mathbb{Z}}.$$

Puisque l'ensemble non errant  $\Omega(\phi_p)$  de  $\phi_p$  est fermé et contient les points périodiques de  $\phi_p$ , nous avons  $\Omega(\phi_p) = \mathbb{R}/\mathbb{Z}$ , et le système dynamique topologique  $(\mathbb{R}/\mathbb{Z}, \phi_p)$  est non errant.

Pour tout  $x \in \mathbb{R}/\mathbb{Z}$ , sa préimage totale  $\bigcup_{k \in \mathbb{N}} \phi_p^{-k}(x)$ , qui vaut

$$\bigcup_{k \in \mathbb{N}} \phi_p^{-k}(x) = \left\{ \frac{x+k}{p^n} \pmod{\mathbb{Z}} : n \in \mathbb{N}, k \in \mathbb{Z} \right\},$$

est dense dans  $\mathbb{R}/\mathbb{Z}$ . Donc pour tous les ouverts non vides  $U$  et  $V$ , en prenant  $x \in V$  (c'est possible car  $V$  est non vide), la propriété de densité ci-dessus et le fait que  $U$  est non vide montrent qu'il existe  $n \in \mathbb{N}$  tel que  $U \cap \phi_p^{-n}(x)$  soit non vide. En particulier,  $U \cap \phi_p^{-n}(V)$  est non vide. Par conséquent,  $\phi_p$  est positivement transitif.

Pour tous les  $x \in \mathbb{R}/\mathbb{Z}$  et  $n \in \mathbb{N}$ , la préimage  $\phi_p^{-n}(x)$  de  $x$  par  $\phi_p^n$  est

$$\phi_p^{-n}(x) = \left\{ \frac{x+k}{p^n} \pmod{\mathbb{Z}} : k \in \mathbb{Z} \right\}$$

et tout point de  $\mathbb{R}/\mathbb{Z}$  est à distance au plus  $\frac{1}{2p^n}$  de l'un des points de  $\phi_p^{-n}(x)$ . Donc pour tous les ouverts  $U$  et  $V$  non vides, soit  $N \in \mathbb{N}$  tel que  $U$  contienne un intervalle ouvert de longueur  $\frac{2}{p^N}$ . Alors pour tout  $n \geq N$ , l'intersection  $U \cap \phi_p^{-n}(V)$  est non vide. Donc  $\phi_p$  est topologiquement mélangeant.

(2) L'application  $\Theta : \Sigma_+ = \{0, \dots, p-1\}^{\mathbb{N}} \rightarrow \mathbb{R}/\mathbb{Z}$  est clairement bien définie et continue, <sup>149</sup> et elle semi-conjugué bien le décalage  $\sigma_+$  et  $\phi_p$ . La surjectivité de  $\Theta$  vient de l'existence du développement  $p$ -adique de tout élément de  $[0, 1[$ . Puisque  $\Theta$  est continue et envoie point périodique sur point périodique, et puisque les points périodiques d'un système de Bernoulli sont denses, nous avons

$$\mathbb{R}/\mathbb{Z} = \Theta(\{0, \dots, p-1\}^{\mathbb{N}}) = \Theta(\overline{\text{Per}(\sigma_+)}) \subset \overline{\Theta(\text{Per}(\sigma_+))} \subset \overline{\text{Per}(\phi_p)} \subset \mathbb{R}/\mathbb{Z},$$

donc  $\text{Per}(\phi_p)$  est dense.

Puisque  $\Theta$  envoie l'ensemble  $\omega$ -limite de tout  $x \in \Sigma_+$  dans l'ensemble  $\omega$ -limite de  $\Theta(x)$ , puisqu'il existe un élément  $x \in \Sigma_+$  tel que  $\omega_{\sigma_+}(x) = \Sigma_+$  (voir l'exemple (2) suivant l'exercice E.46), nous avons donc  $\omega_{\phi_p}(\Theta(x)) = \mathbb{R}/\mathbb{Z}$ . Donc  $\phi_p$  est positivement transitif.

Puisqu'être topologiquement mélangeant est préservé par passage à un facteur (voir l'exercice E.48) et puisqu'un décalage de Bernoulli est topologiquement mélangeant (voir l'exemple (2) suivant l'exercice E.46), ceci remonte que  $\phi_p$  est topologiquement mélangeant.

(3) Il est immédiat que l'opérateur  $L$  est bien défini, linéaire et continu, de norme au plus 1, car  $\|Lf\|_{\infty} \leq \frac{1}{p} \sum_{i=1}^p \|f\|_{\infty} = \|f\|_{\infty}$  pour tout  $f \in C^0(\mathbb{T}^1)$ .

<sup>149</sup>. Si  $x = (x_i)_{i \in \mathbb{N}}$  et  $y = (y_i)_{i \in \mathbb{N}}$  vérifient  $d(x, y) < e^{-N}$  pour la distance définie par la formule (4), alors  $|\Theta(x) - \Theta(y)| \leq \frac{1}{p^{N-1}}$  par sommation de série géométrique.

Montrons la formule (61). Sur chacun des intervalles  $]\frac{k}{p}, \frac{k+1}{p}[$  pour  $k = 0, \dots, p-1$ , l'application  $x \mapsto px$  est un homéomorphisme sur l'intervalle  $]k, k+1[$ . Donc par changements de variables, en identifiant les fonctions sur  $\mathbb{T}^1$  aux fonctions 1-périodiques sur  $\mathbb{R}$ , nous avons

$$\begin{aligned} \int_{\mathbb{T}^1} (f \circ \phi_p) g \, d\lambda &= \sum_{k=0}^{p-1} \int_{\frac{k}{p}}^{\frac{k+1}{p}} f(px) g(x) \, dx = \sum_{k=0}^{p-1} \frac{1}{p} \int_k^{k+1} f(x) g\left(\frac{x}{p}\right) \, dx \\ &= \int_0^1 f(x) \frac{1}{p} \sum_{k=0}^{p-1} g\left(\frac{x+k}{p}\right) \, dx = \int_{\mathbb{T}^1} f(Lg) \, d\lambda. \end{aligned}$$

En prenant pour  $g$  la fonction constante égale à 1, comme  $Lg$  est aussi la fonction constante égale à 1, cette formule dit que l'application  $\phi_p$  préserve la mesure de Lebesgue.

Montrons la formule (62). Soit  $f \in C^1(\mathbb{T}^1)$ . Rappelons que si  $s \leq s'$ , alors nous avons  $f(s') - f(s) = \int_s^{s'} f'(u) \, du$ . Par récurrence, pour tout  $x \in [0, 1[$ , nous avons, en identifiant les fonctions sur  $\mathbb{T}^1$  aux fonctions 1-périodiques sur  $\mathbb{R}$ ,

$$\begin{aligned} \left| L^n f(x) - \int_{\mathbb{T}^1} f \, d\lambda \right| &= \left| \frac{1}{p^n} \left( \sum_{k=0}^{p^n-1} f\left(\frac{x+k}{p^n}\right) \right) - \int_{\frac{x}{p^n}}^{\frac{x}{p^n}+1} f(t) \, dt \right| \\ &= \left| \sum_{k=0}^{p^n-1} \int_{\frac{x+k}{p^n}}^{\frac{x+k+1}{p^n}} \left( f\left(\frac{x+k}{p^n}\right) - f(t) \right) \, dt \right| \\ &\leq \sum_{k=0}^{p^n-1} \int_{\frac{x+k}{p^n}}^{\frac{x+k+1}{p^n}} \int_{\frac{x+k}{p^n}}^t |f'(u)| \, du \, dt \\ &\leq \sum_{k=0}^{p^n-1} \int_{\frac{x+k}{p^n}}^{\frac{x+k+1}{p^n}} \int_{\frac{x+k}{p^n}}^{\frac{x+k+1}{p^n}} |f'(u)| \, du \, dt = \frac{1}{p^n} \int_0^1 |f'(u)| \, du. \end{aligned}$$

Le fait qu'être exponentiellement mélangeant implique être mélangeant a été vu en toute généralité dans la partie 3.2. Montrons donc la formule (63).

Soient  $f, g \in C^1(\mathbb{T}^1)$ . Par les formules (61) et (62), nous avons

$$\begin{aligned} \left| \int_{\mathbb{T}^1} (f \circ \phi_p^n) g \, d\lambda - \left( \int_{\mathbb{T}^1} f \, d\lambda \right) \left( \int_{\mathbb{T}^1} g \, d\lambda \right) \right| &= \left| \int_{\mathbb{T}^1} f(L^n g) \, d\lambda - \left( \int_{\mathbb{T}^1} f \, d\lambda \right) \left( \int_{\mathbb{T}^1} g \, d\lambda \right) \right| \\ &\leq \int_{\mathbb{T}^1} |f| \left| L^n g - \int_{\mathbb{T}^1} g \, d\lambda \right| \, d\lambda \leq \frac{1}{p^n} \|f\|_\infty \|g'\|_\infty \leq \frac{1}{p^n} \|f\|_{1,\infty} \|g\|_{1,\infty}. \end{aligned}$$

**Correction de l'exercice E.48.** • Par la formule (1), nous avons  $h(\text{Per}(\phi)) \subset \text{Per}(\phi')$  et puisque  $h$  est continue, nous avons  $h(\overline{A}) \subset \overline{h(A)}$  pour toute partie  $A$  de  $X$ . Puisque  $h$  est surjective, si  $\text{Per}(\phi)$  est dense, nous avons donc

$$X' = h(X) = h(\overline{\text{Per}(\phi)}) \subset \overline{h(\text{Per}(\phi))} \subset \overline{\text{Per}(\phi')} \subset X'.$$

D'où  $\text{Per}(\phi')$  est dense.

• Si  $x$  est un point positivement récurrent de  $\phi$ , alors  $x \in \omega(x)$ . Donc par la formule (59), nous avons

$$h(x) \in h(\omega(x)) \subset \omega(h(x)),$$

d'où  $h(x)$  est positivement récurrent.

- Nous savons par la formule (60) que  $h^{-1}(\Omega(\phi'))$  contient  $\Omega(\phi)$ . Donc si  $\phi$  n'a pas de point errant, alors, puisque  $h$  est surjective,  $\Omega(\phi') \supset h(\Omega(\phi)) = h(X) = X'$ . D'où  $\phi'$  n'a pas de point errant.

- Si  $(X, \phi)$  est positivement transitif (respectivement transitif dans le cas inversible), alors  $(X', \phi')$  l'est, c'est l'exercice E.45. Voici un autre argument lorsque  $X$  et  $X'$  sont des espaces de Baire à base dénombrable d'ouverts. Si  $(X, \phi)$  est positivement transitif, soit  $x$  un point de  $X$  dont l'ensemble  $\omega$ -limite est égal à tout  $X$ . Par la formule (59), l'ensemble  $\omega$ -limite  $\omega(h(x))$  contient  $h(\omega(x)) = h(X)$ . Par la surjectivité de  $h$ , nous avons donc  $\omega(h(x)) = X$ .

- Supposons que  $(X, \phi)$  soit topologiquement mélangeant. Pour tous les ouverts non vides  $U$  et  $V$  de  $X'$ , puisque  $h^{-1}(U)$  et  $h^{-1}(V)$  sont des ouverts non vides de  $X$ , il existe un temps  $T \geq 0$  tel que, pour tout  $t \geq T$ , l'intersection  $h^{-1}(U) \cap \phi^{-t}(h^{-1}(V))$  soit non vide. Comme l'image d'une intersection est contenue dans l'intersection des images, comme

$$\phi^{-t}(h^{-1}(V)) = (h \circ \phi^t)^{-1}(V) = ((\phi')^t \circ h)^{-1}(V) = h^{-1}((\phi')^{-t}(V)),$$

l'intersection  $U \cap (\phi')^{-t}(V)$ , qui contient  $h(h^{-1}(U) \cap h^{-1}((\phi')^{-t}(V)))$ , est non vide. Donc  $(X', \phi')$  est topologiquement mélangeant.

- Supposons que  $(X, \phi)$  soit positivement minimal. Pour tout  $x' \in X'$ , puisque  $h$  est surjective, soit  $x \in X$  tel que  $h(x) = x'$ . Par la formule (59), l'ensemble  $\omega$ -limite  $\omega(x') = \omega(h(x))$  contient  $h(\omega(x)) = h(X) = X'$ , ces dernières égalités étant vérifiées par la minimalité positive de  $\phi$  et la surjectivité de  $h$ . Donc  $(X, \phi)$  est positivement minimal.

De même, si  $(X, \phi)$  est minimal, pour tout  $x' \in X'$ , soit  $x \in X$  tel que  $h(x) = x'$ . En utilisant la surjectivité de  $h$  pour la première égalité, la densité de l'orbite de  $x$  pour la seconde égalité, la continuité de  $h$  pour la première inclusion, et le fait que  $h$  envoie orbite dans orbite pour la seconde inclusion (voir la formule (3)), nous avons

$$X' = h(X) = h(\overline{\mathcal{O}(x)}) \subset \overline{h(\mathcal{O}(x))} \subset \overline{\mathcal{O}(h(x))} = \overline{\mathcal{O}(x')}.$$

D'où  $(X', \phi')$  est minimal.

**Correction de l'exercice E.49.** Nous ne traitons que le cas du temps discret. Soit  $\mathcal{V}$  une base dénombrable d'ouverts non vides de  $X$ , et pour tout  $V \in \mathcal{V}$ , soit  $V' = \bigcup_{n \in \mathbb{N}} \phi^{-n}(V)$ . Alors  $V'$  est ouvert, car  $\phi$  est continue.

Montrons qu'il est dense, ce qui conclut car par le théorème de Baire (valable pour les espaces métrisables séparables complets), l'intersection dénombrable d'ouverts denses  $\bigcap_{V \in \mathcal{V}} \bigcup_{n \in \mathbb{N}} \phi^{-n}(V)$  est non vide, et tout point de cet ensemble a une orbite positive dense, car rencontrant tout élément d'une base d'ouverts.

Soit  $U$  un ouvert non vide, montrons qu'il rencontre  $V'$ , ce qui conclut. Puisqu'il existe un point d'orbite dense dans  $X$  par l'exercice E.46, il existe  $n \geq 0$  tel que  $U \cap \phi^{-n}(V)$  ou  $U \cap \phi^n(V)$  est non vide. Le premier cas donne le résultat. Dans le second cas, comme tout ouvert non vide contient une infinité de points de toute partie dense par l'hypothèse sur la topologie de  $X$ , il existe  $m > n$  et un point  $x$  de  $U \cap \phi^n(V)$  tel que  $\phi^{-m}(x) \in U \cap \phi^n(V)$ . Mézalor le point  $y = \phi^{-m}(x)$  appartient à  $U$  et  $y \in \phi^{-m}(\phi^n(V)) = \phi^{-(m-n)}(V)$ , ce qui conclut car  $m > n$ .

**Correction de l'exercice E.50.** Soient  $U$  et  $V$  des ouverts non vides. Puisque la mesure  $\mu$  est de support total, nous avons  $\mu(U), \mu(V) \neq 0$ . Si  $(X, \mu, \phi)$  est mélangeant, alors  $\mu(U \cap \phi^{-t}(V))$  converge vers  $\mu(U)\mu(V) > 0$  quand  $t \rightarrow +\infty$ . Donc il existe  $T$  tel que si  $t \geq T$ , alors  $\mu(U \cap \phi^{-t}(V)) > 0$ . En particulier  $U \cap \phi^{-t}(V)$  est non vide, et donc  $(X, \phi)$  est topologiquement mélangeant.

Les rotations du cercle ne sont pas topologiquement mélangeantes, comme vu dans l'exemple (1) de la partie 7.4, donc ne sont pas mélangeantes pour la mesure de Lebesgue du cercle.

**Correction de l'exercice E.51.** (1) Puisque la suite  $(\phi^n)_{n \in \mathbb{N}}$  est équicontinue et comme  $X$  est compact, pour tout  $f \in C^0(X; \mathbb{C})$ , la suite  $(S_n f)_{n \in \mathbb{N}}$  est bornée et équicontinue. Par le théorème d'Arzela-Ascoli (voir par exemple la note de bas de page 147), elle admet donc sous-suite qui converge vers une application  $\tilde{f} \in C^0(X; \mathbb{C})$  (pour la topologie de la convergence uniforme sur  $C^0(X; \mathbb{C})$ ). Celle-ci est continue et  $\phi$ -invariante. Comme  $\phi$  est topologiquement transitive, et puisque  $\tilde{f}$  est constante sur toute orbite, l'application  $\tilde{f}$  est donc constante. Soit  $\mu$  une mesure  $\phi$ -invariante (qui existe par le théorème de Krylov-Bogolyubov, voir la proposition 0.6). Cette constante  $\tilde{f}$  est égale à  $\mu(f)$  car  $\mu(S_n f) = \mu(f)$  et par convergence uniforme. En particulier, elle ne dépend pas de la sous-suite extraite, donc la suite  $(S_n f)_{n \in \mathbb{N}}$  converge vers une constante ne dépendant que de  $f$ . Le résultat découle alors de la proposition 4.2 (3).

(2) Supposons tout d'abord que  $G$  soit métrisable, et fixons  $d_0$  une distance sur  $G$  induisant sa topologie. Il suffit alors de munir  $G$  d'une distance invariante par toutes les translations à gauche, et d'appliquer la question (1). Soit  $\mu$  une mesure de Haar à gauche sur  $G$ . Alors  $d : (x, y) \mapsto \int_G d_0(gx, gy) d\mu(g)$  est une telle distance.

Lorsque  $G$  n'est pas métrisable, il est possible d'étendre la question (1) au cas des espaces uniformes (voir par exemple [Bou]), mais voici un autre argument. Notons  $g \in G$  l'élément de  $G$  tel que  $\tau = \lambda_g$  soit la translation à gauche par  $g$ . Par hypothèse, il existe  $x \in G$  tel que la suite  $(g^n x)_{n \in \mathbb{N}}$  soit dense. Par la continuité de la translation à droite par  $x^{-1}$ , la suite  $(g^n)_{n \in \mathbb{N}}$  est donc dense dans  $G$ . Soient  $\mu$  une mesure de probabilité  $\tau$ -invariante sur  $G$  et  $h \in G$ . Il existe une sous-suite  $(g^{n_k})_{k \in \mathbb{N}}$  qui converge vers  $h$ . Donc  $\lambda_{g^{n_k}}$  converge vers  $\lambda_h$  dans  $C^0(X; X)$ . Donc par la continuité pour la topologie uniforme de  $C^0(X; X)$  et celle de la convergence vague sur  $\text{Prob}(X)$  de l'application  $f \mapsto f_* \mu$  image réciproque d'une mesure fixée, et par l'invariance de  $\mu$  par  $\tau = \lambda_g$ , nous avons

$$(\lambda_h)_* \mu = \lim_{k \rightarrow +\infty} (\lambda_{g^{n_k}})_* \mu = \mu .$$

Donc  $\mu$  est en fait invariante par toutes les translations, donc coïncide avec la mesure de Haar normalisée de  $G$ . Le résultat découle alors de la proposition 4.2 (2)

(3) Nous laissons le soin au lecteur de montrer que l'application  $\Theta$  de  $\llbracket 0, p-1 \rrbracket^{\mathbb{N}}$  dans  $\mathbb{Z}_p$  définie par  $(a_i)_{i \in \mathbb{N}} \mapsto \sum_{i \in \mathbb{N}} a_i p^i$  (en prenant la limite, qui existe, pour la distance  $p$ -adique étendue par complétion à  $\mathbb{Z}_p$ ) est un homéomorphisme de l'espace de Cantor  $\llbracket 0, p-1 \rrbracket^{\mathbb{N}}$  sur  $\mathbb{Z}_p$ , qui est un isomorphisme de groupes si nous munissons  $\llbracket 0, p-1 \rrbracket^{\mathbb{N}}$  de l'addition terme à terme avec possible retenue ajoutée au terme suivant, et qui envoie les suites presque nulles sur l'ensemble  $\mathbb{N}$ , par le développement  $p$ -adique des entiers : tout élément  $n \in \mathbb{N}$  non nul s'écrit de manière unique  $n = \sum_{i=0}^k a_i p^i$  où  $k \in \mathbb{N}$  et  $a_i \in \llbracket 0, p-1 \rrbracket$  avec  $a_k \neq 0$ .

Par transport de structure, les images par  $\Theta$  des cylindres initiaux  $[a_0, \dots, a_k]$  de l'espace produit  $\llbracket 0, p-1 \rrbracket^{\mathbb{N}}$  pour  $k \in \mathbb{N}$  et  $a_0, \dots, a_k \in \llbracket 0, p-1 \rrbracket$  forment une base d'ouverts

dans  $\mathbb{Z}_p$ . Si  $n = \sum_{i=0}^{k'} a'_i p^i$  avec  $k' \geq k$  et  $a'_0 = a_0, \dots, a'_k = a_k$ , alors il est immédiat de montrer que  $\phi^n(0) \in \Theta([a_0, \dots, a_k])$ . Donc l'orbite de 0 par la transformation de l'odomètre est dense. Il suffit alors d'appliquer la question (2).

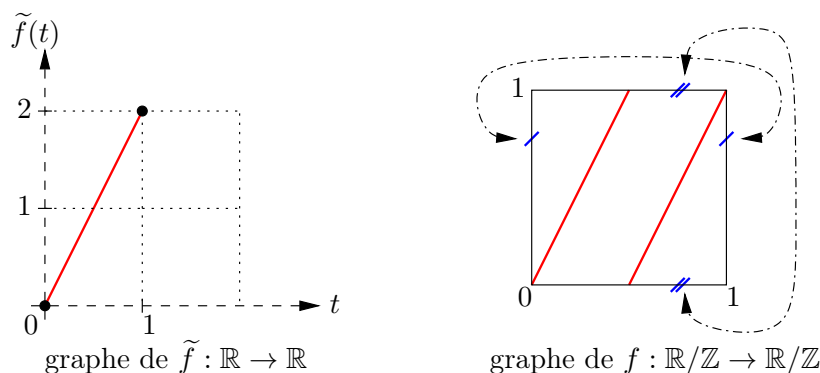
## 8 Dynamique des homéomorphismes du cercle

Rappelons (voir par exemple [Pau4, Pau8] pour des rappels de topologie algébrique élémentaire) que l'application lisse  $\rho$  de  $\mathbb{R}$  dans  $\mathbb{S}_1$  définie par  $\theta \mapsto e^{2i\pi\theta}$  est un revêtement universel de  $\mathbb{S}_1$ , de groupe de revêtement  $\mathbb{Z}$ . Il induit par passage au quotient un  $C^\infty$ -difféomorphisme  $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{S}_1$  par lequel nous identifions ces deux variétés différentielles lisses.<sup>150</sup> Alors  $\rho : \mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$  est la projection canonique  $t \mapsto \dot{t} = t \bmod \mathbb{Z}$ . Par le théorème de relèvement, pour toute application continue  $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$ , il existe une application  $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$  continue, unique modulo translation entière, appelée un *relevé* de  $f$ , telle que le diagramme suivant commute

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\tilde{f}} & \mathbb{R} \\ \rho \downarrow & & \downarrow \rho \\ \mathbb{R}/\mathbb{Z} & \xrightarrow{f} & \mathbb{R}/\mathbb{Z}, \end{array}$$

c'est-à-dire telle que  $\rho \circ \tilde{f} = f \circ \rho$ . Le nombre  $d = \tilde{f}(1) - \tilde{f}(0)$  ne dépend pas du choix du relevé  $\tilde{f}$ . Il est appelé le *degré* de  $f$  et noté  $\deg(f)$ , et nous avons  $\tilde{f}(x+1) = \tilde{f}(x) + d$  pour tout  $x \in \mathbb{R}$ . Il vérifie  $\deg(\text{id}) = 1$  et  $\deg(f \circ g) = \deg(f) \deg(g)$  pour toute application continue  $g : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$ .

**Exemple.** Pour tout  $d \in \mathbb{N} - \{0\}$ , le degré d'un revêtement de degré  $d$  du cercle est égal à  $d$ . En particulier, l'application de doublement de l'angle,<sup>151</sup> qui est l'application  $f : x \mapsto 2x$  de  $\mathbb{R}/\mathbb{Z}$  dans  $\mathbb{R}/\mathbb{Z}$ , est de degré 2. Pour tout  $k \in \mathbb{Z}$ , il existe de même au moins une application continue  $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  telle que  $\deg(f) = k$ .



Si  $f$  est un homéomorphisme du cercle, alors  $\deg(f) = \pm 1$  et  $\tilde{f}$  est un homéomorphisme de  $\mathbb{R}$ . De plus,  $f$  préserve l'orientation du cercle si et seulement si  $\deg(f) = 1$ , et si

150. En munissant la variété lisse quotient  $\mathbb{R}/\mathbb{Z}$  de sa structure de groupe additif abélien quotient (donc de  $\mathbb{Z}$ -module, de sorte que pour tout  $x \in \mathbb{R}/\mathbb{Z}$ , l'élément  $2x = x + x$  est bien défini), et la sous-variété lisse  $\mathbb{S}_1$  de sa structure de groupe multiplicatif, ce  $C^\infty$ -difféomorphisme  $\mathbb{R}/\mathbb{Z} \rightarrow \mathbb{S}_1$  est aussi un isomorphisme de groupes.

151. Voir les exercices E.47 et E.53 pour des compléments.

et seulement si l'application  $\tilde{f}$  est strictement croissante. Par exemple, les relevés de la rotation  $z \mapsto e^{i\theta} z$  du cercle  $\mathbb{S}_1$ , d'angle  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ , sont les applications  $t \mapsto t + \alpha + k$  où  $\alpha = \frac{\theta}{2\pi} \bmod \mathbb{Z}$  et  $k \in \mathbb{Z}$ . Le réel  $\alpha$  est bien défini modulo  $\mathbb{Z}$ , et cette rotation, notée  $R_\alpha$  lorsque nous la considérons comme une application de  $\mathbb{R}/\mathbb{Z}$  dans  $\mathbb{R}/\mathbb{Z}$ , est rationnelle si et seulement si  $\alpha \in \mathbb{Q}$ , et irrationnelle si et seulement si  $\alpha \in \mathbb{R} - \mathbb{Q}$ . En rappelant que  $\mathbb{R}/\mathbb{Z}$  est un groupe additif, nous appellerons encore  $\alpha \in \mathbb{R}/\mathbb{Z}$  l'angle de la rotation

$$R_\alpha : x \mapsto x + \alpha .$$

## 8.1 Nombre de rotation d'un homéomorphisme du cercle

**Proposition 8.1.** *Soit  $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  un homéomorphisme préservant l'orientation. Alors pour tout  $x$  dans  $\mathbb{R}$  et pour tout relevé  $\tilde{f}$  de  $f$ , la limite*

$$\tau(\tilde{f}) = \lim_{n \rightarrow +\infty} \frac{1}{n} (\tilde{f}^n(x) - x)$$

existe dans  $\mathbb{R}$ , est indépendante du choix du point  $x$  dans  $\mathbb{R}$ , et

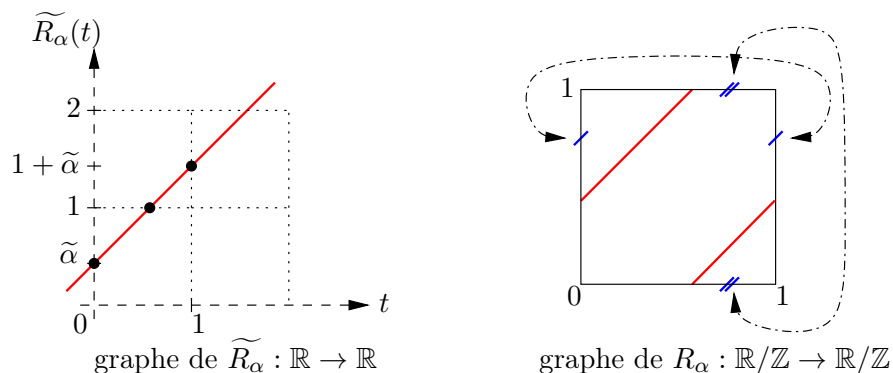
$$\sup_{m \in \mathbb{N}} |\tilde{f}^m(x) - m\tau(\tilde{f}) - x| \leq 2 . \quad (64)$$

Elle est appelée le nombre de translation de  $\tilde{f}$ . Sa classe modulo  $\mathbb{Z}$  est indépendante du choix du relevé  $\tilde{f}$  de  $f$ . Elle est notée  $\rho(f) \in \mathbb{R}/\mathbb{Z}$ , et appelée le nombre de rotation de  $f$ .

**Exemple.** Pour tout  $\alpha$  dans le groupe additif  $\mathbb{R}/\mathbb{Z}$ , le nombre de rotation de la rotation  $R_\alpha : x \mapsto x + \alpha$  est égal à l'angle de cette rotation :

$$\rho(R_\alpha) = \alpha \in \mathbb{R}/\mathbb{Z} .$$

En effet, en notant  $\tilde{\alpha}$  un relevé de  $\alpha$  dans  $\mathbb{R}$ , l'application  $\tilde{R}_\alpha : x \mapsto x + \tilde{\alpha}$  est un relevé de  $R_\alpha$  (les autres sont de la forme  $x \mapsto x + \tilde{\alpha} + k$  avec  $k \in \mathbb{Z}$ ), et donc  $\tilde{R}_\alpha^n(x) = x + n\tilde{\alpha}$  pour tout  $n \in \mathbb{N}$ , ce qui montre le résultat.



**Démonstration.** Montrons l'existence de la limite. Rappelons le lemme suivant bien connu des suites presque sous-additives (voir par exemple [Cal] pour une introduction aux quasi-morphismes de groupes).



**Lemme 8.2.** Soient  $c \in \mathbb{R}$  et  $(a_n)_{n \in \mathbb{N}}$  une suite réelle telle que

$$\forall n, m \in \mathbb{N}, \quad a_{n+m} \leq a_n + a_m + c.$$

Alors la limite  $\lim_{n \rightarrow +\infty} \frac{a_n}{n}$  existe dans  $\mathbb{R} \cup \{-\infty\}$ . Si  $\ell \in \mathbb{R}$ , alors

$$\sup_{m \in \mathbb{N}} (m\ell - a_m) \leq c.$$

**Démonstration.** Soit  $\ell = \inf_{n \geq 1} \frac{a_n}{n} \in \mathbb{R} \cup \{-\infty\}$ . Supposons tout d'abord  $\ell > -\infty$ . Alors la suite positive  $b_n = a_n - \ell n$  est encore presque sous-additive pour la même constante  $c$ , et pour tout  $\epsilon > 0$ , il existe  $n_0 \geq 1$  tel que  $\frac{b_{n_0}}{n_0} \leq \epsilon$  et  $\frac{1}{n_0} \leq \epsilon$ . Pour tout  $n \geq n_0$ , écrivons  $n = qn_0 + r$  avec  $0 \leq r < n_0$ . Par presque sous-additivité et récurrence, nous avons alors

$$\frac{b_n}{n} \leq \frac{qb_{n_0} + rb_1 + (q+r-1)c}{n} \leq \frac{b_{n_0}}{n_0} + \frac{1}{n_0} + \frac{n_0(b_1 + c)}{n}.$$

Donc  $0 < \frac{b_n}{n} \leq 3\epsilon$  pour tout  $n$  assez grand, ce qui montre que  $\frac{a_n}{n}$  converge vers  $\ell$ .

Si  $\ell = -\infty$ , pour tout  $A > 0$ , il existe  $n_0 \geq 1$  tel que  $\frac{a_{n_0}}{n_0} \leq -A - 2$ . Comme ci-dessus, pour tout  $n \geq n_0$ , si  $n = qn_0 + r$  avec  $0 \leq r < n_0$  alors

$$\frac{a_n}{n} \leq \frac{qa_{n_0} + ra_1 + (q+r-1)c}{n} \leq \frac{a_{n_0}}{n_0} + \frac{1}{n_0} + \frac{n_0(a_1 + c)}{n}.$$

Donc  $\frac{a_n}{n} \leq -A$  pour tout  $n$  assez grand, ce qui montre que  $\frac{a_n}{n}$  converge vers  $-\infty$ .

Par presque sous-additivité et récurrence, pour tous les entiers  $m, n \in \mathbb{N}$ , nous avons  $a_{mn} \leq na_m + (n-1)c$ . Donc

$$m \frac{a_{mn}}{mn} - a_m = \frac{a_{mn} - na_m}{n} \leq c.$$

En faisant tendre  $n$  vers  $+\infty$ , le résultat en découle.  $\square$

Pour tout entier  $n \in \mathbb{N}$ , posons  $a_n = \tilde{f}^n(x) - x$  et  $k_n = \lfloor a_n \rfloor$ . Pour tout entier  $m \in \mathbb{N}$ , puisque  $f^m$  est un homéomorphisme préservant l'orientation, remarquons que  $\tilde{f}^m(x+k) = f^m(x) + k$  pour tout  $k \in \mathbb{Z}$ , que  $|f^m(y) - f^m(z)| \leq 1$  si  $|y - z| \leq 1$  et que  $\tilde{f}^n(x) - (x+k_n) \in [0, 1]$ . Alors

$$\begin{aligned} & |a_{m+n} - a_m - a_n| \\ &= |(\tilde{f}^m(\tilde{f}^n(x)) - \tilde{f}^m(x+k_n) + \tilde{f}^m(x) + k_n - x) - (\tilde{f}^m(x) - x) - (\tilde{f}^n(x) - x)| \\ &\leq |\tilde{f}^m(\tilde{f}^n(x)) - \tilde{f}^m(x+k_n)| + |k_n + x - \tilde{f}^n(x)| \leq 2. \end{aligned}$$

Si  $n \geq 1$ , nous avons, par somme télescopique et puisque la fonction  $y \mapsto \tilde{f}(y) - y$  est 1-périodique,

$$\frac{a_n}{n} = \frac{1}{n} \sum_{i=0}^{n-1} (\tilde{f}(\tilde{f}^i(x)) - \tilde{f}^i(x)) \geq \inf_{y \in [0,1]} (\tilde{f}(y) - y).$$

Donc la suite  $(\frac{a_n}{n})_{n \in \mathbb{N} - \{0\}}$ , qui converge dans  $\mathbb{R} \cup \{-\infty\}$  par le lemme 8.2 et qui est minorée par la continuité de  $\tilde{f}$  et la compacité de  $[0, 1]$ , admet une limite dans  $\mathbb{R}$ . Ceci

montre l'existence du nombre de translation  $\tau(\tilde{f})$ , ainsi que la formule (64), en appliquant le lemme 8.2 à  $(a_n)_{n \in \mathbb{N}}$  et  $(-a_n)_{n \in \mathbb{N}}$  avec  $c = 2$ .

Soit  $n \in \mathbb{N} - \{0\}$ . Puisque la fonction  $y \mapsto \tilde{f}^n(y) - y$  est 1-périodique, le nombre de translation  $\tau(\tilde{f})$  ne dépend que de  $x$  modulo  $\mathbb{Z}$ . Si  $x, y \in [0, 1]$ , alors  $|\tilde{f}^n(x) - \tilde{f}^n(y)| \leq 1$ , donc

$$\left| \frac{1}{n} (\tilde{f}^n(x) - x) - \frac{1}{n} (\tilde{f}^n(y) - y) \right| \leq \frac{1}{n} |\tilde{f}^n(x) - \tilde{f}^n(y)| + \frac{1}{n} |x - y| \leq \frac{2}{n}$$

Donc le nombre de translation  $\tau(\tilde{f})$  ne dépend pas de  $x$ .

Si  $F$  est un autre relevé de  $f$ , et si  $k \in \mathbb{Z}$  est tel que  $F(t) - \tilde{f}(t) = k$  pour tout  $t \in \mathbb{R}$ , alors

$$F^n(x) - x = \tilde{f}^n(x) - x + kn$$

pour tout  $n \in \mathbb{N}$ . Donc le nombre de rotation  $\rho(f)$  dans  $\mathbb{R}/\mathbb{Z}$  ne dépend pas du choix du relevé de  $f$ .  $\square$

Le résultat suivant montre que le nombre de rotation est un invariant de conjugaison, et décrit la situation quand le nombre de rotation est rationnel.

**Proposition 8.3.** *Soit  $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  un homéomorphisme préservant l'orientation.*

- (1) *Nous avons  $\rho(f^m) = m \rho(f)$  pour tout  $m \in \mathbb{N}$ .*
- (2) *Le nombre de rotation  $\rho(f)$  est rationnel si et seulement si  $f$  admet un point périodique.*
- (3) *Si  $h : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  est un homéomorphisme préservant l'orientation, alors*

$$\rho(h \circ f \circ h^{-1}) = \rho(f) .$$

**Démonstration.** (1) Notons que  $\tilde{f}^m$  est un relevé de  $f^m$ . Nous avons

$$\rho(f) = \lim_{n \rightarrow +\infty} \frac{1}{nm} (\tilde{f}^{nm}(x) - x) \pmod{\mathbb{Z}} = \frac{1}{m} \rho(f^m) .$$

(2) Si  $f$  admet un point périodique de période  $q$  et si  $x$  est un relevé de ce point périodique dans  $\mathbb{R}$ , alors il existe un entier  $p \in \mathbb{Z}$  tel que  $\tilde{f}^q(x) = x + p$ . Donc par récurrence, pour tout  $n \in \mathbb{N}$ ,

$$\frac{1}{nq} (\tilde{f}^{nq}(x) - x) = \frac{np}{nq} = \frac{p}{q}$$

Donc  $\rho(f) = \frac{p}{q} \pmod{\mathbb{Z}}$  est rationnel.

Réciproquement, si  $\rho(f) = \frac{p}{q} \in \mathbb{Q}/\mathbb{Z}$ , alors  $\rho(f^q) = 0$  par l'assertion (1). Il suffit donc de montrer que si  $\rho(f) = 0$ , alors  $f$  a un point fixe. Supposons par contraposition que  $f$  n'a pas de point fixe. Soit  $\tilde{f}$  un relevé de  $f$  tel que  $\tilde{f}(0) \in [0, 1[$ . Pour tout  $x \in \mathbb{R}$ , le réel  $\tilde{f}(x) - x$  n'est pas entier (sinon l'image de  $x$  dans  $\mathbb{R}/\mathbb{Z}$  serait fixe). Par le théorème des valeurs intermédiaires, nous avons donc que  $\tilde{f}(x) - x \in ]0, 1[$  pour tout  $x \in \mathbb{R}$ . Par continuité et 1-périodicité, il existe donc  $\delta > 0$  tel que  $\tilde{f}(x) - x \in ]\delta, 1 - \delta[$  pour tout  $x \in \mathbb{R}$ . Donc, pour tout  $n \in \mathbb{N} - \{0\}$ ,

$$\frac{1}{n} (\tilde{f}^n(0) - 0) = \frac{1}{n} \sum_{i=0}^{n-1} (\tilde{f}^{i+1}(0) - \tilde{f}^i(0)) \in ]\delta, 1 - \delta[ .$$

Ceci montre que  $\rho(f) \neq 0$ .

(3) Soient  $\tilde{f}$  et  $\tilde{h}$  des relevés de  $f$  et  $h$  respectivement, tels que  $\tilde{h}(0) \in [0, 1[$ . Notons que  $\tilde{h} \circ \tilde{f} \circ \tilde{h}^{-1}$  est un relevé de  $h \circ f \circ h^{-1}$ . Puisque  $\tilde{h}$  est un homéomorphisme croissant de  $\mathbb{R}$  qui se projette en un homéomorphisme modulo  $\mathbb{Z}$ , pour tout  $x \in [0, 1[$ , nous avons

$$0 - 1 < \tilde{h}(x) - x \leq \tilde{h}(x) \leq \tilde{h}(1) \leq 2,$$

et par 1-périodicité, nous avons  $|\tilde{h}(x) - x| \leq 2$  pour tout  $x \in \mathbb{R}$ . De même, nous avons  $|\tilde{h}^{-1}(x) - x| \leq 2$  pour tout  $x \in \mathbb{R}$ . Soit  $n \in \mathbb{N} - \{0\}$ . Pour tous les  $x, y \in \mathbb{R}$ , si  $|y - x| \leq 2$ , alors, en posant  $z = \frac{x+y}{2}$ , de sorte que  $|z - x| \leq 1$  et  $|z - y| \leq 1$ , nous avons

$$|\tilde{f}^n(y) - \tilde{f}^n(x)| \leq |\tilde{f}^n(y) - \tilde{f}^n(z)| + |\tilde{f}^n(z) - \tilde{f}^n(x)| \leq 2.$$

D'où

$$\begin{aligned} & |\tilde{h} \circ \tilde{f}^n \circ \tilde{h}^{-1}(x) - \tilde{f}^n(x)| \\ & \leq |\tilde{h}(\tilde{f}^n \circ \tilde{h}^{-1}(x)) - \tilde{f}^n \circ \tilde{h}^{-1}(x)| + |\tilde{f}^n(\tilde{h}^{-1}(x)) - \tilde{f}^n(x)| \leq 4. \end{aligned}$$

Par conséquent, nous avons  $\frac{1}{n} |(\tilde{h} \circ \tilde{f} \circ \tilde{h}^{-1})^n(x) - \tilde{f}^n(x)| \leq \frac{4}{n}$ , ce qui montre que  $\rho(h \circ f \circ h^{-1}) = \rho(f)$ .  $\square$

**Exercice E.52.** Soit  $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  un homéomorphisme préservant l'orientation, de nombre de rotation  $\alpha$  rationnel.

(1) Montrer que toutes les orbites périodiques ont la même période.

(2) Montrer que toutes les orbites périodiques ont le même ordre cyclique sur le cercle que toute orbite de la rotation  $R_\alpha$ .

(3) Si le nombre de rotation de  $f$  est  $p/q$  avec  $p$  et  $q$  deux entiers premiers entre eux, montrer que  $f$  est topologiquement conjugué à la rotation  $R_{p/q}$  si et seulement s'il existe un relevé  $\tilde{f}$  de  $f$  tel que  $\tilde{f}^q$  soit l'application  $t \mapsto t + p$ .

## 8.2 Théorie de Poincaré-Denjoy<sup>144</sup> des homéomorphismes du cercle

**Théorème 8.4** (Denjoy). Soit  $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  un difféomorphisme de classe  $C^2$  préservant l'orientation, de nombre de rotation  $\alpha$  irrationnel. Alors  $f$  est topologiquement conjugué à la rotation  $R_\alpha$ .

**Remarques.** (1) Nous verrons que  $f$  n'est pas forcément  $C^2$ -conjugué à  $R_\alpha$ . Le théorème reste vrai si  $f$  est de classe  $C^1$  et si l'application<sup>152</sup>  $\log f' : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}$  est à variation bornée<sup>153</sup> (voir [KH, §12.1]).

152. Si  $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  est un difféomorphisme de classe  $C^1$  préservant l'orientation, alors la dérivée  $\tilde{f}' : \mathbb{R} \rightarrow \mathbb{R}$  d'un relevé  $\tilde{f}$  de  $f$  ne dépend pas du choix de ce relevé. Elle est 1-périodique et strictement positive, donc définit une application  $f' : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}$  strictement positive.

153. Une fonction  $g : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}$  est à variation bornée si la borne supérieure, sur toutes les suites finies  $(I_k)_{1 \leq k \leq n}$  d'arcs de cercle  $I_k$ , d'extrémités  $x_k, y_k$ , d'intérieurs deux à deux disjoints,

$$\text{Var}(g) = \sup_{(I_k)_{1 \leq k \leq n}} \sum_{k=1}^n |g(y_k) - g(x_k)|$$

est finie. Notons que si  $g$  est de classe  $C^1$ , alors  $g$  est à variation bornée et  $\text{Var}(g) = \int_0^1 |g'(t)| dt$ .

(2) Montrons qu'une conjugaison  $h : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  telle que  $h \circ f = R_\alpha \circ h$  est uniquement déterminée modulo postcomposition par une rotation : si  $h^*$  est une autre conjugaison, alors il existe  $\beta \in \mathbb{R}/\mathbb{Z}$  tel que  $h^* = R_\beta \circ h$ . Lorsque le choix de la conjugaison n'est pas important, nous parlerons donc par abus de "la" conjugaison.

En effet, puisque deux rotations commutent,  $R_\beta \circ h$  est encore une conjugaison entre  $f$  et  $R_\alpha$ . Quitte à postcomposer  $h$  et  $h^*$  par des rotations, nous pouvons supposer que  $h$  et  $h^*$  ont deux relevés qui conjuguent un relevé  $\tilde{f}$  de  $f$  et coïncident en 0. Ces deux relevés coïncident donc sur la partie  $\{f^n(0) + m : m, n \in \mathbb{Z}\}$  de  $\mathbb{R}$ , qui est dense par conjugaison et car  $\{n\alpha + m : n, m \in \mathbb{Z}\}$  l'est (par l'irrationalité de  $\alpha$ ). Ils coïncident donc partout par continuité de  $h$  et  $h^*$ .

La première étape de la démonstration du théorème 8.4, due à Poincaré, est de montrer que  $f$  est topologiquement semi-conjugue à  $R_\alpha$ , donc que le système dynamique topologique  $(\mathbb{R}/\mathbb{Z}, R_\alpha)$  est un facteur du système dynamique topologique  $(\mathbb{R}/\mathbb{Z}, f)$ , avec un complément disant que si  $f$  admet une orbite dense, alors les deux systèmes sont topologiquement conjugués.

**Théorème 8.5** (Poincaré). *Soit  $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  un homéomorphisme préservant l'orientation du cercle, de nombre de rotation  $\alpha$  irrationnel. Il existe une application continue monotone<sup>154</sup> surjective  $h : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  telle que  $h \circ f = R_\alpha \circ h$ . Si  $f$  est transitif,<sup>155</sup> alors  $h$  est un homéomorphisme, qui conjugue  $f$  et  $R_\alpha$ .*

Nous donnons deux démonstrations, au fond semblables.

**Démonstration 1.** (Voir par exemple [Ghy].) Soit  $\tilde{f}$  un relevé de  $f$ . L'application  $\tilde{h} : \mathbb{R} \rightarrow \mathbb{R}$  définie par

$$\tilde{h} : x \mapsto \sup_{m \in \mathbb{N}} (\tilde{f}^m(x) - m \tau(\tilde{f}))$$

est à valeurs finies par la formule (64) et vérifie

- (i)  $\tilde{h}(x + 1) = \tilde{h}(x) + 1$  pour tout  $x \in \mathbb{R}$ ,
- (ii)  $\tilde{h}(\tilde{f}(x)) = \tilde{h}(x) + \tau(\tilde{f})$  pour tout  $x \in \mathbb{R}$ .

Par (i) et puisque  $\alpha = \rho(f) = \tau(\tilde{f}) \bmod \mathbb{Z}$ , elle induit donc par passage au quotient une application  $h : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  telle que  $h \circ f = R_\alpha \circ h$ . L'application  $\tilde{h}$  est croissante, car  $\tilde{f}$  l'est, et semi-continue inférieurement car borne supérieure de fonctions continues. En particulier, elle est continue à gauche. Donc son image est le complémentaire d'une union dénombrable (invariante par la translation  $t \mapsto t + \tau(\tilde{f})$  par (ii)) d'intervalles semi-ouverts disjoints (les « sauts » de  $\tilde{h}$ ), et  $\tilde{h}$  est continue si et seulement si cette union est vide. Or cette union est la préimage par la projection canonique  $\mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$  du complémentaire de l'image de  $h$ . Puisque  $\alpha$  est irrationnel et par (ii), l'image de  $h$  est dense, donc  $\tilde{h}$  est continue.

L'application continue croissante  $\tilde{h}$  est strictement croissante en dehors d'une union dénombrable (invariante par l'application  $\tilde{f}$  par (ii)) d'intervalles fermés maximaux d'intérieur non vide deux à deux disjoints, sur chacun desquels  $\tilde{h}$  est constante (les « plats »

154. Une application continue du cercle dans le cercle est *monotone* si au moins un (donc tout) relevé de  $\mathbb{R}$  dans  $\mathbb{R}$  de cette application est monotone.

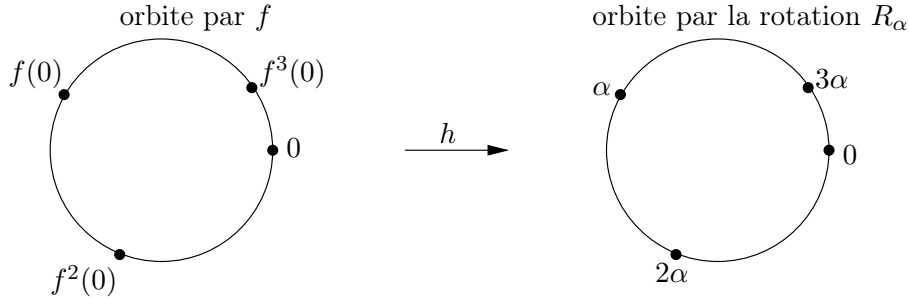
155. c'est-à-dire si le système dynamique topologique  $(\mathbb{R}/\mathbb{Z}, f)$  est (topologiquement) transitif, voir la partie 7.4

de  $\tilde{h}$ ). Notons  $U$  l'image dans  $\mathbb{R}/\mathbb{Z}$  de la réunion des intérieurs de ces intervalles, et  $\omega_f$  le fermé complémentaire, qui est invariant par  $f$ .

Si  $f$  est transitive, c'est-à-dire admet une orbite dense par l'exercice E.46, alors  $U$  est vide, et donc  $\tilde{h}$  est injective. Par conséquent,  $h$  est une bijection continue de  $\mathbb{R}/\mathbb{Z}$ , donc un homéomorphisme par compacité, qui conjugue  $f$  et  $R_\alpha$ . En particulier, toute orbite de  $f$  est dense, et le système dynamique topologique  $(\mathbb{R}/\mathbb{Z}, f)$  est minimal.

Si  $f$  n'est pas transitive, alors  $\omega_f$  est un espace de Cantor (fermé, sans point isolé et d'intérieur vide du cercle). La préimage d'un point de  $\mathbb{R}/\mathbb{Z}$  par la restriction à  $\omega_f$  de  $\tilde{h}$  est constituée de 1 ou 2 points, ces deux points étant alors les extrémités d'une composante connexe de  $U$ . Les points de  $U$  sont errants pour  $f$ , et l'ensemble non errant de  $f$  est  $\omega_f$ . Toute orbite de  $\omega_f$  est dense dans  $\omega_f$ , qui est donc l'unique minimal de  $f$ , appelé un *minimal exceptionnel*.

**Démonstration 2.** La remarque clef pour cette deuxième démonstration du théorème 8.5 (donnée par la première assertion du lemme suivant) est que les orbites sur le cercle de l'homéomorphisme  $f$  sont cycliquement ordonnées exactement comme celles de la rotation  $R_\alpha$ .



Soit  $f$  comme dans l'énoncé du théorème 8.5. Nous supposons que nous avons fixé un relevé  $\tilde{\alpha} \in \mathbb{R} - \mathbb{Q}$  dans  $\mathbb{R}$  du nombre de rotation  $\alpha \in (\mathbb{R} - \mathbb{Q})/\mathbb{Z}$  de  $f$ , et un relevé  $\tilde{f}$  de  $f$  de sorte que  $\tilde{\alpha} = \lim_{k \rightarrow +\infty} \frac{\tilde{f}^k(0)}{k}$ .

**Lemme 8.6.** *Nous avons les propriétés suivantes.*

(1) *Pour tous les  $x \in \mathbb{R}$  et  $n, m, n', m' \in \mathbb{Z}$ , nous avons*

$$n\tilde{\alpha} + m < n'\tilde{\alpha} + m' \text{ si et seulement si } \tilde{f}^n(x) + m < \tilde{f}^{n'}(x) + m'$$

(2) *Pour tous les  $x \in \mathbb{R}/\mathbb{Z}$  et  $m < n$  dans  $\mathbb{Z}$ , nous avons  $f^n(x) \neq f^m(x)$  et si  $I$  est l'un des deux arcs de cercles d'extrémités  $f^n(x)$  et  $f^m(x)$ , alors toute orbite positive de  $f$  rencontre  $I$ .*

(3) *L'ensemble  $\omega$ -limite d'un point de  $\mathbb{R}/\mathbb{Z}$  est indépendant de ce point, et noté  $\omega_f$ , et ou bien  $\omega_f = \mathbb{R}/\mathbb{Z}$  (auquel cas  $f$  est transitif), ou bien le fermé  $\omega_f$  est sans point isolé, d'intérieur vide (auquel cas  $f$  n'est pas transitif), et est l'unique minimal de  $f$ .*

La dernière assertion du lemme ci-dessus dit que  $f$  admet un unique minimal, et ou bien toute orbite de  $f$  est dense (et donc le cercle tout entier est le minimal de  $f$ ), ou bien l'unique minimal de  $f$  est un ensemble de Cantor. Un tel minimal est dit *exceptionnel*.

**Démonstration.** (1) Montrons que la quantité  $(\tilde{f}^n(x) + m) - (\tilde{f}^{n'}(x) + m')$  ne change pas de signe quand  $x$  varie dans  $\mathbb{R}$ , et donc que l'inégalité  $\tilde{f}^n(x) + m < \tilde{f}^{n'}(x) + m'$  est

indépendante de  $x$ . Sinon, par le théorème des valeurs intermédiaires, il existerait  $y \in \mathbb{R}$  tel que  $\tilde{f}^n(y) + m = \tilde{f}^{n'}(y) + m'$ , et l'image de  $y$  dans le cercle  $\mathbb{R}/\mathbb{Z}$  serait un point périodique de  $f$ , contredisant la proposition 8.3 (2).

Supposons que

$$\tilde{f}^n(0) + m < \tilde{f}^{n'}(0) + m'. \quad (65)$$

Montrons que  $n\tilde{\alpha} + m < n'\tilde{\alpha} + m'$ . Supposons par exemple que  $n \geq n'$ , le résultat se démontre de manière analogue sinon (prendre  $k \in -\mathbb{N}$ ,  $k \neq 0$ , et la somme de  $i = -1$  à  $i = k$  dans ce qui suit). En posant  $y = \tilde{f}^{n'}(0)$ , l'inégalité (65) équivaut à  $\tilde{f}^{n-n'}(y) - y < m' - m$ . Comme ci-dessus, cette inégalité est vérifiée pour tout  $y \in \mathbb{R}$ , donc pour tout  $k \in \mathbb{N}$ ,  $k \neq 0$ , nous avons

$$\tilde{f}^{k(n-n')}(0) = \sum_{i=1}^k (\tilde{f}^{i(n-n')}(0) - \tilde{f}^{(i-1)(n-n')}(0)) < k(m' - m).$$

Si  $n = n'$ , ceci implique que  $m' > m$ , ce qui conclut. Sinon  $n > n'$  et

$$\tilde{\alpha} = \lim_{k \rightarrow +\infty} \frac{\tilde{f}^{k(n-n')}(0)}{k(n-n')} < \lim_{k \rightarrow +\infty} \frac{k(m' - m)}{k(n-n')} = \frac{m' - m}{n - n'},$$

ce qui conclut.

(2) Pour tout  $k \in \mathbb{N}$ , posons  $I_k = f^{-k(n-m)}(I)$ . Remarquons que  $f^{n-m}$  envoie  $f^m(x)$  sur  $f^n(x)$ , et donc les intervalles  $I_0 = I$  et  $I_1 = f^{-(n-m)}(I)$  ont une extrémité en commun (le point  $f^m(x)$ ). Par conséquent, pour tout  $k \in \mathbb{N}$ , les intervalles  $I_k$  et  $I_{k+1}$  ont une extrémité en commun (le point  $f^{(k+1)m-kn}(x)$ ). Supposons par l'absurde que nous ayons  $\mathbb{R}/\mathbb{Z} \neq \bigcup_{k \in \mathbb{N}} I_k$ . Alors les extrémités de  $I_k$  convergent vers un point  $z \in \mathbb{R}/\mathbb{Z}$ , qui est fixé par  $f^{n-m}$ . Donc  $z$  est périodique, ce qui contredit la proposition 8.3 (2).

(3) Soient  $x, y \in \mathbb{R}/\mathbb{Z}$ , montrons que les ensembles  $\omega$ -limites de  $x$  et  $y$  coïncident. Soient  $z \in \omega(x)$  et  $(n_k)_{k \in \mathbb{N}}$  une suite strictement croissante dans  $\mathbb{N}$  telle que  $\lim_{k \rightarrow +\infty} f^{n_k}(x) = z$ . Pour tout  $k$  assez grand, notons  $I_k$  celui des deux arcs de cercle d'extrémités  $f^{n_k}(x)$  et  $f^{n_{k+1}}(x)$  dont la longueur est la plus petite, et donc tend vers 0 quand  $k \rightarrow +\infty$ . Par l'assertion (2), l'orbite positive de  $y$  rencontre  $I_k$  pour tout  $k$ , donc  $z \in \omega(y)$ . Par conséquent, nous avons  $\omega(x) \subset \omega(y)$ , avec égalité par symétrie.

Par la proposition 7.4, le système dynamique topologique  $(X, f)$  admet au moins un minimal. Puisque tout minimal de  $f$  contient l'ensemble  $\omega$ -limite de chacun de ses points, le fermé  $\omega_f$ , qui est non vide par compacité et invariant par  $f$  d'après l'exercice E.42, est l'unique minimal de  $f$ . Comme tout minimal, il est ou bien égal à tout l'espace, ou bien d'intérieur vide (voir la remarque (2) de la partie 7.4).

Pour tout  $x$  dans  $\omega_f$ , puisque  $x \in \omega_f = \omega(x)$ , il existe une suite strictement croissante  $(n_k)_{k \in \mathbb{N}}$  dans  $\mathbb{N}$  telle que  $\lim_{k \rightarrow +\infty} f^{n_k}(x) = x$ . Si  $x$  est isolé dans  $\omega_f$ , alors  $f^{n_k}(x) = x$  pour  $k$  assez grand, donc  $x$  est périodique, ce qui contredit la proposition 8.3 (2).  $\square$

Revenons à la démonstration du théorème 8.5. Pour construire l'application  $h$ , nous commençons par la construire sur l'orbite par  $f$  du point  $0 \in \mathbb{R}/\mathbb{Z}$  en l'envoyant en préservant le temps sur l'orbite par  $R_\alpha$  de 0. Nous l'étendrons alors par continuité à l'adhérence de son orbite, puis de manière localement constante sur son complémentaire, s'il existe.

Plus précisément, soit  $A = \{\tilde{f}^n(0) + m : m, n \in \mathbb{Z}\}$  le relevé total de l'orbite de  $0 \in \mathbb{R}/\mathbb{Z}$ . Nous définissons  $\tilde{h} : A \rightarrow \mathbb{R}$  comme l'unique application strictement croissante

envoyant  $\tilde{f}^n(0) + m$  sur  $n\tilde{\alpha} + m$ , ce qui est possible par le lemme 8.6 (1). En notant  $\tilde{R}_\alpha : \mathbb{R} \rightarrow \mathbb{R}$  l'application  $t \mapsto t + \tilde{\alpha}$ , par construction, nous avons, en tout point de  $A$ ,

$$\tilde{h} \circ \tilde{f} = \tilde{R}_\alpha \circ \tilde{h}.$$

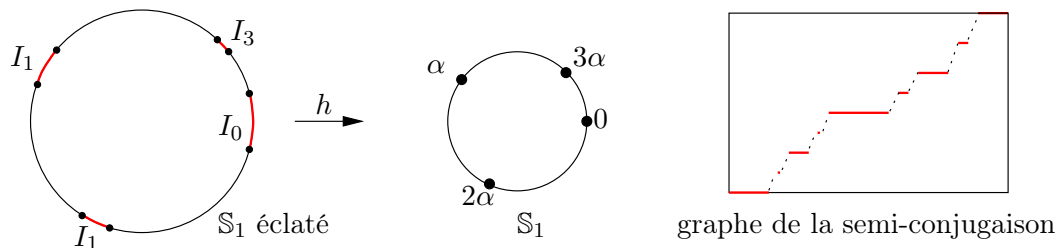
Montrons que  $\tilde{h}$  s'étend continûment à l'adhérence  $\bar{A}$  de  $A$  dans  $\mathbb{R}$ . Puisque  $\tilde{h}$  est croissante, les limites à droites et à gauche de  $\tilde{h}$  existent en tout point de  $\bar{A}$ , et si elles étaient différentes, le complémentaire de l'adhérence de l'ensemble  $\tilde{h}(A) = \{n\tilde{\alpha} + m : m, n \in \mathbb{Z}\}$  serait non vide, ce qui contredirait le fait qu'il est dense dans  $\mathbb{R}$ , puisque  $\tilde{\alpha}$  est irrationnel. Par densité de  $\tilde{h}(A)$ , l'application  $\tilde{h} : \bar{A} \rightarrow \mathbb{R}$  est de plus surjective.

Puisque  $\tilde{h} : \bar{A} \rightarrow \mathbb{R}$  est croissante et surjective, pour toute composante connexe  $I$  de  $\mathbb{R} - \bar{A}$  (qui est un intervalle ouvert et borné par invariance de  $\bar{A}$  par translations entières), les valeurs de  $\tilde{h}$  sur les extrémités de  $I$  sont égales, et nous étendons donc continûment  $\tilde{h}$  à  $\mathbb{R}$  en envoyant chaque tel intervalle  $I$  de manière constante sur son adhérence.

Par construction, nous avons  $\tilde{h}(x+1) = \tilde{h}(x) + 1$  sur  $A$ , donc sur  $\bar{A}$  par extension continue, puis sur  $\mathbb{R}$  par la naturalité de l'extension ci-dessus. Donc  $\tilde{h}$  induit par passage au quotient une application continue monotone surjective  $h : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  qui semi-conjuge  $f$  et  $R_\alpha$ .

Lorsque  $f$  est transitive, nous avons  $\bar{A} = \mathbb{R}$  et donc  $h$  est continue bijective, donc un homéomorphisme par compacité.  $\square$

Il est assez facile de construire des exemples d'homéomorphismes préservant l'orientation du cercle, de nombre de rotation irrationnel prescrit, pour lesquels la semi-conjugaison construite ci-dessus n'est pas une conjugaison.



Si  $\alpha \in \mathbb{R}/\mathbb{Z}$  est irrationnel, fixons  $(\ell_n)_{n \in \mathbb{Z}}$  une suite sommable de réels strictement positifs. Partons de la rotation irrationnelle  $R_\alpha$ , et éclatons une <sup>156</sup> orbite, au sens suivant. Enlevons du cercle l'orbite complète de  $0 \in \mathbb{R}/\mathbb{Z}$  par  $R_\alpha$ , et remplaçons le point  $R_\alpha^n(0)$  par un intervalle  $I_n$  de longueur  $\ell_n$  de manière à préserver l'ordre. Considérons la topologie qui fait qu'une suite de points du complémentaire de l'orbite qui s'accumulait par la gauche sur  $R_\alpha^n(0)$  s'accumule maintenant sur l'extrémité gauche de  $I_n$  et une qui s'accumulait par la droite sur  $R_\alpha^n(0)$  s'accumule maintenant sur l'extrémité droite de  $I_n$ . Puisque la suite réelle  $(\ell_n)_{n \in \mathbb{Z}}$  est sommable, nous obtenons ainsi un nouveau cercle topologique.

Considérons alors la transformation  $f$  sur ce cercle qui coïncide avec  $R_\alpha$  en dehors de  $\bigcup_{n \in \mathbb{Z}} I_n$ , en demandant qu'elle envoie l'intervalle  $I_n$  de manière croissante affine sur l'intervalle  $I_{n+1}$ . L'application  $f$  est un homéomorphisme préservant l'orientation du cercle, non transitif (les points intérieurs des intervalles  $I_k$  sont errants). Le nombre de rotation de  $f$  est  $\alpha$ . L'unique minimal de  $f$  est l'ensemble de Cantor adhérence du complémentaire

156. Il serait aussi possible d'éclater un ensemble dénombrable d'orbites.

de  $\bigcup_{n \in \mathbb{Z}} I_n$ . La semi-conjugaison  $h$  entre  $f$  et  $R_\alpha$  est constante sur chaque intervalle  $I_n$ , et son graphe est donc du type de l'*escalier du diable de Lebesgue*, voir la figure ci-dessus à droite.

En travaillant un peu plus, il est même possible de construire des exemples de classe  $C^1$ , et nous renvoyons par exemple à [KH, §12.1] pour une démonstration du résultat suivant, et à [Her] pour des compléments.

**Proposition 8.7** (Denjoy). *Pour tous les  $\rho \in (\mathbb{R} - \mathbb{Q})/\mathbb{Z}$  et  $\alpha \in ]0, 1[$ , il existe un  $C^1$ -difféomorphisme non transitif du cercle, de dérivée  $\alpha$ -höldérienne (voir la note de bas de page 89), de nombre de rotation  $\rho$ .  $\square$*

Par contre, le but du théorème de Denjoy 8.4 est de montrer que de tels exemples ne sont pas possibles en classe  $C^2$ .

**Démonstration du théorème 8.4.** L'argument principal est d'utiliser la régularité supplémentaire pour obtenir des résultats de non distorsion permettant d'exclure le cas non transitif (c'est-à-dire l'existence d'un minimal exceptionnel), ce qui permet d'appliquer le théorème 8.5.

Supposons par l'absurde que  $f$  ne soit pas topologiquement conjugué à la rotation  $R_\alpha$ . Alors par le théorème 8.5, l'unique minimal  $\omega_f$  de  $f$  est un ensemble de Cantor, et nous notons  $I$  une composante connexe de son complémentaire. Elle est disjointe de toutes ses images par les applications  $f^n$  pour  $n \in \mathbb{Z} - \{0\}$ .

Rappelons que pour tout  $n \in \mathbb{Z}$ , l'homéomorphisme  $f^n$  préserve l'orientation, donc que sa dérivée est minorée par une constante strictement positive, par compacité. Notons

$$V = \int_0^1 |(\log f')'(t)| dt = \sup_{(I'_k)_{1 \leq k \leq n}} \sum_{k=1}^n |\log f'(y_k) - \log f'(x_k)| \quad (66)$$

la variation totale de  $\log f'$ , où la borne supérieure est prise sur toutes les suites finies  $(I'_k)_{1 \leq k \leq n}$  d'intervalles  $I'_k$  de  $\mathbb{R}/\mathbb{Z}$ , d'extrémités  $x_k, y_k$ , d'intérieurs deux à deux disjoints.

**Lemme 8.8.** *Il existe une infinité d'entiers  $n \in \mathbb{N}$  tels que pour tout  $x \in I$ ,*

$$e^{-V} \leq (f^n)'(x) (f^{-n})'(x) \leq e^V .$$

Ce lemme de non distorsion permet de conclure, car pour tout  $n$  dans la partie infinie de  $\mathbb{N}$  donnée par le lemme, nous avons

$$\begin{aligned} \int_{f^n(I)} dx + \int_{f^{-n}(I)} dx &= \int_I (f^n)'(x) + (f^{-n})'(x) dx \\ &\geq \int_I \sqrt{(f^n)'(x) (f^{-n})'(x)} dx \geq e^{-V/2} \int_I dx . \end{aligned}$$

Donc la somme des longueurs des intervalles  $f^n(I)$  pour  $n \in \mathbb{Z}$  est infinie, ce qui contredit le fait qu'ils sont deux à deux disjoints dans le cercle.

**Démonstration du lemme 8.8.** Puisque  $I$  s'envoie sur un point par la semi-conjugaison  $h$  entre  $f$  et la rotation  $R_\alpha$ , et puisque l'orbite de  $I$  est ordonnée sur le cercle comme celle de la rotation irrationnelle  $R_\alpha$  et est dense, il existe une partie infinie  $\mathcal{N}$  de  $\mathbb{N}$  telle que pour



tous les  $n \in \mathcal{N}$  et  $x \in I$ , si  $I'_n = ]x, f^{-n}(x)[$ , alors les intervalles  $f^k(I'_n)$  pour  $0 \leq k < n$  sont deux à deux disjoints.

Donc pour tous les  $n \in \mathcal{N}$  et  $x \in I$ , en posant  $y = f^{-n}(x)$ , par la formule (66) pour la première inégalité et par la formule de dérivation des fonctions composées pour la dernière égalité, nous avons

$$\begin{aligned} V &\geq \sum_{k=0}^{n-1} \left| \log f'(f^k(x)) - \log f'(f^k(y)) \right| \geq \left| \sum_{k=0}^{n-1} \log f'(f^k(x)) - \log f'(f^k(y)) \right| \\ &= \left| \log \frac{\prod_{k=0}^{n-1} f'(f^k(x))}{\prod_{k=0}^{n-1} f'(f^k(y))} \right| = \left| \log \frac{(f^n)'(x)}{(f^n)'(y)} \right|, \end{aligned}$$

ce qui montre le lemme 8.8 car  $y = f^{-n}(x)$ , en utilisant la formule de dérivation des inverses.  $\square$

Ceci conclut la démonstration du théorème 8.4.  $\square$

### 8.3 Problème de $C^r$ -conjugaison et petits diviseurs

Soit  $r$  une régularité dans  $\mathbb{N} \cup \{\infty, \omega\}$  (muni de l'ordre usuel sur  $\mathbb{N}$  complété par  $n < \infty < \omega$  pour tout  $n \in \mathbb{N}$ ).<sup>157</sup> Le problème de la régularité d'une conjugaison  $h$  entre un  $C^r$ -difféomorphisme  $f$  du cercle de nombre de rotation  $\alpha$  irrationnel et la rotation  $R_\alpha$  d'angle  $\alpha$  (lorsqu'elle existe, par exemple si  $r \geq 2$ ) est une question naturelle, mais qui n'a pas de réponse élémentaire. Elle a nécessité des développements profonds, par Arnold,<sup>158</sup> Moser<sup>158</sup>, Herman<sup>158</sup> et Yoccoz<sup>158</sup> entre autres. Je recommande fortement de regarder une présentation historique en français sur YouTube

<https://www.youtube.com/watch?v=9Cfq80ffjDw>

des travaux de Herman et Yoccoz, en particulier la présentation de Chenciner, et le numéro spécial de la Gazette des mathématiciens consacré à Yoccoz. Notons que la régularité de  $h$  ne dépend pas du choix de la conjugaison, par la remarque (2) suivant l'énoncé du théorème 8.4.

Cette régularité n'est absolument pas automatique : même si  $f$  est très régulière,  $h$  ne l'est pas forcément. En fait Arnold (voir par exemple [KH, §12.5]) a construit de nombreux exemples de difféomorphismes analytiques réels  $f$  du cercle, de nombre de rotation irrationnel, dont la conjugaison  $h$  n'est pas de classe  $C^1$  (et même n'est pas absolument

<sup>157</sup>. La régularité  $C^\omega$  est la régularité analytique réelle. Étant donné un intervalle  $I$  de  $\mathbb{R}$  et un espace de Banach réel  $E$ , une application  $g : I \rightarrow E$  est *analytique réelle* si elle est localement développable en série entière normalement convergente : pour tout  $t_0 \in I$ , il existe une suite  $(c_n)_{n \in \mathbb{N}}$  dans  $E$  et un voisinage  $U$  de  $t_0$  dans  $I$  tels que pour tout  $t \in U$ , la série  $\sum_{n \in \mathbb{N}} (t - t_0)^n c_n$  soit normalement convergente dans  $E$ , de somme égale à  $g(t)$ . Nous renvoyons par exemple à [Car] pour des informations. Une application continue du cercle dans le cercle est *analytique réelle* si au moins un (donc tout) relevé l'est.



158. Vladimir Arnold  
1937-2010



Jürgen Moser  
1929-1999



Michaël Herman  
1942-2000



Jean-Christophe Yoccoz  
1958-2016



Klaus Roth  
1925-2015

continue par rapport à la mesure de Lebesgue  $\text{Leb}_{\mathbb{S}^1}$  (c'est-à-dire que les mesures  $h_* \text{Leb}_{\mathbb{S}^1}$  et  $\text{Leb}_{\mathbb{S}^1}$  ne sont pas absolument continues l'une par rapport à l'autre)).

Pour éviter une perte de régularité pour la conjugaison, il se trouve être nécessaire d'introduire des conditions arithmétiques sur les nombres de rotation, disant qu'ils sont mal approchables par des rationnels.

Pour tout  $\beta \geq 0$ , nous dirons qu'un irrationnel  $\alpha \in \mathbb{R} - \mathbb{Q}$  vérifie une *condition diophantienne d'ordre  $\beta$*  s'il existe une constante  $\gamma > 0$  telle que

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{\gamma}{q^{2+\beta}}$$

pour tout rationnel  $p/q$ . Un irrationnel  $\alpha$  est dit *diophantien* s'il vérifie une condition diophantienne d'ordre  $\beta$  pour un certain  $\beta \geq 0$ . Il est dit *de Liouville* s'il n'est pas diophantien. Notons  $C_\beta$  l'ensemble des irrationnels qui vérifient une condition diophantienne d'ordre  $\beta$ , et  $\text{Dio}$  l'ensemble des nombres diophantiens. Un théorème de Roth<sup>27</sup> (lui ayant valu une médaille Fields en 1958) dit que presque tout réel (pour la mesure de Lebesgue) est *de type Roth*, c'est-à-dire appartient à  $\bigcap_{\beta > 0} C_\beta$ .

Un résultat d'Herman dit que la condition sur un irrationnel  $\alpha$  d'être diophantien est nécessaire à l'existence d'une conjugaison  $C^\infty$  pour tout difféomorphisme de nombre de rotation  $\alpha$ . Le résultat suivant de Yoccoz, pour lequel nous renvoyons à [Yoc1], dit en particulier que l'ensemble des irrationnels  $\alpha$  tels que tout difféomorphisme  $C^\infty$  préservant l'orientation du cercle de nombre de rotation  $\alpha$  est  $C^\infty$ -conjugué à la rotation  $R_\alpha$  est exactement l'ensemble  $\text{Dio}$  des irrationnels diophantiens. En régularité moindre que  $C^\infty$ , il y a une perte de régularité de la conjugaison. Rappelons que pour tous les  $\alpha \in ]0, 1[$  et  $r \in \mathbb{N}$ , un  $C^{r+\alpha}$ -difféomorphisme du cercle est un  $C^r$ -difféomorphisme du cercle  $f$  dont la dérivée  $f^{(r)}$  d'ordre  $r$  est  $\alpha$ -höldérienne.

**Théorème 8.9** (Yoccoz (1984)). *Soit  $\alpha$  un nombre diophantien d'ordre  $\beta \geq 0$ . Si  $k \in \mathbb{N}$  vérifie  $k > \max\{2, 2\beta + 1\}$ , alors tout  $C^k$ -difféomorphisme préservant l'orientation du cercle, de nombre de rotation  $\alpha$ , est conjugué à la rotation  $R_\alpha$  par un difféomorphisme  $h$  de classe  $C^{k-1-\beta-\epsilon}$  pour tout  $\epsilon > 0$ .  $\square$*

## 8.4 Exercices

**Exercice E.53.** Soit  $(X, d)$  un espace métrique. Une application continue  $f : X \rightarrow X$  est dite *strictement dilatante de rapport  $\lambda > 1$*  s'il existe  $\epsilon > 0$  tel que pour tous les  $x, y \in X$ , si  $d(x, y) \leq \epsilon$  alors

$$d(f(x), f(y)) \geq \lambda d(x, y).$$

Une application continue  $f : X \rightarrow X$  est dite *strictement dilatante* s'il existe  $\lambda > 1$  tel que  $f$  soit strictement dilatante de rapport  $\lambda$ .

Le but de l'exercice est de donner une classification à conjugaison près des applications strictement dilatantes du cercle dans lui-même.<sup>159</sup> Nous considérons le cercle comme l'espace métrique  $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$ , pour la distance  $d$  induite<sup>160</sup> de celle de  $\mathbb{R}$ .

<sup>159.</sup> Nous renvoyons à [Gro] pour la généralisation optimale de ce résultat.

<sup>160.</sup> En notant  $\hat{x} \in \mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$  l'image de  $x \in \mathbb{R}$  par la projection canonique, la distance induite est définie par  $d(\hat{x}, \hat{y}) = \min_{k \in \mathbb{Z}} |x - y + k|$  pour tous les  $x, y \in \mathbb{R}$ . Remarquons que si  $x, y \in \mathbb{R}$  vérifient  $|x - y| < \frac{1}{2}$ , alors  $d(\hat{x}, \hat{y}) = |x - y|$ .

(1) Soit  $p \in \mathbb{Z} - \{-1, 0, 1\}$ . Montrer que l'application  $\phi_p : \mathbb{T}^1 \rightarrow \mathbb{T}^1$  définie par  $x \mapsto px$  est strictement dilatante.

(2) Montrer que le degré  $p$  d'une application strictement dilatante  $f : \mathbb{T}^1 \rightarrow \mathbb{T}^1$  est différent de  $-1, 0, 1$ .

À partir de maintenant, nous fixons une application strictement dilatante  $f : \mathbb{T}^1 \rightarrow \mathbb{T}^1$  de degré  $p \geq 2$ , et nous voulons montrer que  $f$  est topologiquement conjuguée à  $\phi_p$ .

(3) Montrer que  $f$  a un point fixe. Montrer que l'on peut supposer que  $f$  admet un relevé  $\tilde{f}$  (que nous fixons dans la suite) tel que  $\tilde{f}(0) = 0$ . Montrer qu'il existe une suite finie strictement croissante de points  $a_0 = 0 < a_1 < \dots < a_{p-1} < a_p = 1$  tels que  $\tilde{f}(a_i) = i$  pour  $i = 0, \dots, p$ .

(4) Notons  $E$  l'ensemble des applications continues croissantes  $h : [0, 1] \rightarrow [0, 1]$  telles que  $h(0) = 0$  et  $h(1) = 1$ , muni de la distance induite par la norme uniforme

$$d(h_1, h_2) = \max_{t \in [0, 1]} |h_1(t) - h_2(t)|.$$

Montrer que  $E$  est complet. Soit  $L : E \rightarrow E$  l'application  $h \mapsto Lh$  définie en posant, pour tous les  $h \in E$  et  $t \in [0, 1]$ ,

$$Lh : t \mapsto \begin{cases} \frac{1}{p} h(\tilde{f}(t) - i) + \frac{i}{p} \pmod{1} & \text{si } a_i \leq t < a_{i+1} \text{ et } i = 0, \dots, p-1, \\ 1 & \text{si } t = 1. \end{cases}$$

Montrer que  $L$  est une application  $\frac{1}{p}$ -contractante<sup>161</sup> de  $E$  dans  $E$ .

(5) Conclure.

## 8.5 Indications pour la résolution des exercices

**Correction de l'exercice E.52.** Écrivons  $\alpha = \frac{p}{q} \pmod{\mathbb{Z}}$  avec  $p, q \in \mathbb{Z}$  premiers entre eux et  $q > 0$ . Nous pouvons supposer que  $q \geq 2$  (si  $q = 1$ ; alors  $f = \text{id}$ ) et que  $p \in \{0, \dots, q-1\}$ . Notons  $\pi : x \mapsto \dot{x}$  la projection canonique de  $\mathbb{R}$  sur  $\mathbb{R}/\mathbb{Z}$ .

(1) Soit  $x \in \mathbb{R}$  tel que  $\dot{x}$  soit périodique pour  $f$ . Soit  $\tilde{f}$  un relevé de  $f$ . Quitte à le remplacer par un autre relevé, nous pouvons supposer que  $\lim_{n \rightarrow +\infty} \frac{1}{n} (\tilde{f}^n(x) - x) = \frac{p}{q}$ . Puisque  $\dot{x}$  est périodique, il existe donc  $r, s \in \mathbb{Z}$  avec  $r > 0$  tels que  $\tilde{f}^r(x) = x + s$ . Alors

$$\frac{p}{q} = \lim_{n \rightarrow +\infty} \frac{1}{nr} (\tilde{f}^{nr}(x) - x) = \lim_{n \rightarrow +\infty} \frac{ns}{nr} = \frac{s}{r}.$$

Il existe donc  $m \in \mathbb{N} - \{0\}$  tel que  $s = mp$  et  $r = mq$ .

Montrons que  $\tilde{f}^q(x) = x + p$ , ce qui montre l'assertion (1). Si par l'absurde nous avons  $\tilde{f}^q(x) - p > x$ , alors puisque  $\tilde{f}$  est croissante,

$$\tilde{f}^{2q}(x) - 2p = \tilde{f}^q(\tilde{f}^q(x) - p) - p \geq \tilde{f}^q(x) - p > x.$$

Par récurrence,  $\tilde{f}^r(x) - s = \tilde{f}^{mq}(x) - mp > x$ , ce qui contredit la définition de  $r, s$ . Donc  $\tilde{f}^q(x) - p \leq x$  et de même  $\tilde{f}^q(x) - p \geq x$ , ce qui donne le résultat.

<sup>161</sup>. Soit  $c \in [0, 1[$ . Une application  $f : X \rightarrow Y$  entre deux espaces métriques est  $c$ -contractante si pour tous les  $x, y \in X$ , nous avons  $d(f(x), f(y)) \leq c d(x, y)$ .

(2) Puisque  $p$  et  $q$  sont premiers entre eux, l'entier  $p$  est inversible modulo  $q$  et il existe  $k \in \{0, \dots, q-1\}$  tel que  $kp = 1 \pmod q$ . Soit  $x \in \mathbb{R}$  tel que  $\dot{x}$  soit périodique pour  $f$ . Les intervalles

$$[x, \tilde{f}(x)[, [\tilde{f}(x), \tilde{f}^2(x)[, \dots, [\tilde{f}^{q-1}(x), \tilde{f}^q(x)[ \quad (67)$$

sont consécutifs, deux à deux disjoints et de réunion l'intervalle  $[x, x+p[$ . Cet intervalle semi-ouvert de longueur  $p$  contient exactement  $pq$  éléments de

$$A = \mathcal{P}^{-1}(\{\dot{x}, f(\dot{x}), \dots, f^{q-1}(\dot{x})\}) .$$

Puisque  $\tilde{f}$  envoie  $[\tilde{f}^i(x), \tilde{f}^{i+1}(x)[$  sur  $[\tilde{f}^{i+1}(x), \tilde{f}^{i+2}(x)[$  pour  $i = 0, \dots, q-2$  et préserve  $A$ , chacun des  $q$  intervalles de la liste (67) contient exactement  $p$  points de  $A$ . Le point de  $A$  le plus proche de  $x$  dans  $[x, \tilde{f}(x)[$  est de la forme  $\tilde{f}^{q'}(x) - p'$  où  $p' \in \mathbb{Z}$  et  $q' \in \{0, \dots, q-1\}$ . Puisque l'application

$$F : t \mapsto \tilde{f}^{q'}(t) - p'$$

est strictement croissante et préserve  $A$ , et puisque  $[x, \tilde{f}(x)[$  contient exactement  $p$  points de  $A$ , nous avons donc  $F^p(x) = \tilde{f}(x)$ . Par conséquent  $f^{q'p}(\dot{x}) = f(\dot{x})$ , et donc  $q'p = 1 \pmod q$ , d'où  $q' = k$ . Comme l'orbite  $\dot{x}, R_{\frac{p}{q}}(\dot{x}), R_{\frac{2p}{q}}(\dot{x}), \dots, R_{\frac{kp}{q}}(\dot{x})$  de  $\dot{x}$  par la rotation  $R_{\frac{p}{q}}$  est cycliquement ordonnée comme la suite  $\dot{x}, \dot{x} + \frac{kp}{q}, \dot{x} + 2\frac{kp}{q}, \dots, \dot{x} + q\frac{kp}{q}$ , l'assertion (2) en découle.

(3) Si  $f$  est topologiquement conjuguée à la rotation d'angle  $p/q$ , alors  $f^q(x) = x$  pour tout  $x \in \mathbb{R}/\mathbb{Z}$ . Donc si  $\tilde{f}$  est un relevé de  $f$ , pour tout  $x \in \mathbb{R}$ , il existe  $p_x \in \mathbb{Z}$  tel que  $\tilde{f}^q(x) = x + p_x$ , et par continuité l'application  $p_x$  est constante. Par la définition du nombre de rotation, cette constante est égale à  $p$  modulo  $q$ , donc égal à  $p$  quitte à modifier le relevé de  $f$ .

Réciproquement, supposons qu'il existe un relevé  $\tilde{f}$  de  $f$  tel que  $\tilde{f}^q(x) = x + p$  pour tout  $x \in \mathbb{R}$ . Alors toutes les orbites de  $f$  sont périodiques, de même période  $q$  par l'assertion (1), et cycliquement ordonnées comme les orbites de la rotation  $R_{\frac{p}{q}}$  par l'assertion (2). Une conjugaison  $h$  entre  $f$  et  $R_{\frac{p}{q}}$  est obtenue, avec  $F$  la notation introduite dans l'assertion (2), par passage au quotient de l'application  $\tilde{h}$  envoyant linéairement  $[0, F(0)]$  sur l'intervalle  $[0, \frac{1}{q}]$ , puis étendue sur  $[0, 1]$  de manière à ce qu'elle conjugue  $F$  et la translation  $t \mapsto t + \frac{1}{q}$ . Le fait que toutes les orbites aient la même période  $q$  et soient correctement ordonnées montre que  $\tilde{h}$  peut être étendue de manière  $\mathbb{Z}$ -équivariante sur  $\mathbb{R}$ .

**Correction de l'exercice E.53.** (1) Si  $x, y \in \mathbb{R}$  vérifient  $|x - y| < \frac{1}{2|p|}$ , alors

$$d(\phi_p(\dot{x}), \phi_p(\dot{y})) = |px - py| = |p||x - y| = |p|d(\dot{x}, \dot{y}) ,$$

et comme  $|p| \geq 2$ , le résultat en découle.

(2) Soit  $\lambda > 1$  tel que  $f$  soit dilatante de rapport  $\lambda$ . Puisque  $f$  est continue et que l'espace métrique  $\mathbb{T}^1$  est compact, l'application  $f$  est uniformément continue. Si  $\tilde{f}$  est un relevé de  $f$ , alors  $\tilde{f}$  est strictement monotone, car l'application  $(x, y) \mapsto \tilde{f}(x) - \tilde{f}(y)$  a un signe constant sur  $\{(x, y) \in \mathbb{R}^2 : 0 < y - x < \epsilon'\}$  pour  $\epsilon'$  assez petit, donc pour tout  $\epsilon' > 0$ . Par conséquent, pour  $n$  assez grand pour que  $\frac{1}{n}$  soit inférieur au  $\epsilon > 0$  apparaissant dans

la définition d'une application dilatante, nous avons

$$\begin{aligned} |p| &= |\tilde{f}(1) - \tilde{f}(0)| = \left| \sum_{k=1}^n \left( \tilde{f}\left(\frac{k}{n}\right) - \tilde{f}\left(\frac{k-1}{n}\right) \right) \right| = \sum_{k=1}^n \left| \tilde{f}\left(\frac{k}{n}\right) - \tilde{f}\left(\frac{k-1}{n}\right) \right| \\ &\geq \lambda \sum_{k=1}^n \left| \frac{k}{n} - \frac{k-1}{n} \right| = \lambda > 1. \end{aligned}$$

(3) Soient  $\tilde{f}$  un relevé de  $f$  et  $g : x \mapsto \tilde{f}(x) - x$ . Alors

$$g(1) = \tilde{f}(1) - 1 = \tilde{f}(0) + p - 1 \geq \tilde{f}(0) + 1 = g(0) + 1.$$

Donc  $g([0, 1])$  contient un entier. Si  $x \in \mathbb{R}$  vérifie  $g(x) \in \mathbb{Z}$ , alors  $\hat{x}$  est un point fixe de  $f$ . Quitte à conjuguer  $f$  par une rotation (ce qui ne change pas son caractère dilatant), nous pouvons supposer que ce point fixe est 0.

L'application  $\tilde{f}$  est donc un homéomorphisme de  $[0, 1]$  sur  $[0, p]$ , et  $a_1, a_2, \dots, a_{p-1}$  sont les images réciproques de  $1, 2, \dots, p-1$  par cet homéomorphisme.

(4) L'espace  $E$  est un fermé de l'espace de Banach  $C^0([0, 1])$  des applications continues de  $[0, 1]$  dans  $\mathbb{R}$  pour la norme uniforme, donc est complet. Puisque le degré  $p$  de  $f$  est strictement positif, l'application  $\tilde{f}$  est strictement croissante, donc  $i \leq \tilde{f}(t) < i+1$  si  $a_i \leq t < a_{i+1}$ . L'application  $Lh$  est continue sur  $a_i < t < a_{i+1}$ , et continue à droite en  $a_i$ . Elle est aussi continue à gauche, car  $h(1) = 1$  et  $h(0) = 0$ . Comme  $h$  est à valeurs dans  $[0, 1]$ , l'application  $Lh$  l'est aussi. Clairement  $Lh(0) = 0$  et  $Lh(1) = 1$ . Comme  $h$  et  $\tilde{f}$  sont croissantes, l'application  $Lh$  est aussi croissante. Donc  $L$  est bien définie. Pour tous les  $i = 1, \dots, p-1$  et  $h_1, h_2 \in E$ , nous avons

$$\begin{aligned} \max_{t \in [a_i, a_{i+1}[} |L(h_1)(t) - L(h_2)(t)| &= \frac{1}{p} \max_{t \in [a_i, a_{i+1}[} |h_1(f(t) - i) - h_2(f(t) - i)| \\ &\leq \frac{1}{p} \max_{t' \in [0, 1]} |h_1(t') - h_2(t')|, \end{aligned}$$

donc  $L$  est bien  $\frac{1}{p}$ -contractante.

(5) Par le théorème du point fixe <sup>162</sup> des applications strictement contractantes (nous avons

---

162. Ce théorème est le suivant.

**Théorème 8.10. (Théorème du point fixe de Banach)** Soit  $E$  un espace métrique complet non vide et soit  $f : E \rightarrow E$  une application  $c$ -contractante où  $0 < c < 1$ . Alors  $f$  admet un unique point fixe  $\ell$  dans  $E$ , qui vérifie  $d(\ell, x_0) \leq \frac{1}{1-c} d(f(x_0), x_0)$  pour tout  $x_0 \in E$ .

**Démonstration.** Si  $x$  et  $y$  sont des points fixes distincts, alors  $d(x, y) = d(f(x), f(y)) \leq cd(x, y) < d(x, y)$ , une contradiction.

Soit  $x_0$  un point de  $X$ , considérons la suite  $(x_n)_{n \in \mathbb{N}}$  définie par la donnée de  $x_0$  et la relation de récurrence  $x_{n+1} = f(x_n)$ . Montrons que cette suite converge vers un point fixe de  $f$ . Par récurrence, pour tout  $n \in \mathbb{N}$ , nous avons  $d(x_{n+1}, x_n) \leq c^n d(x_1, x_0)$ . Donc par inégalité triangulaire, pour tous les  $n, p \in \mathbb{N}$ ,

$$d(x_{n+p}, x_n) \leq \sum_{k=0}^{p-1} d(x_{n+k+1}, x_{n+k}) \leq \frac{c^n}{1-c} d(x_1, x_0). \quad (68)$$

Par conséquent, la suite  $(x_n)_{n \in \mathbb{N}}$  est de Cauchy dans  $X$ , donc converge vers un point  $x$  de  $X$ . L'application  $f$  étant continue, par passage à la limite quand  $n \rightarrow +\infty$  dans l'équation  $x_{n+1} = f(x_n)$ , nous avons  $x = f(x)$ .

La dernière affirmation du théorème 8.10 découle de la formule (68) en prenant  $n = 0$  et en faisant tendre  $p$  vers  $+\infty$ .

$\frac{1}{p} < 1$ ), soit  $h$  un point fixe de  $L$ . Notons  $\bar{h}$  l'application induite sur  $\mathbb{T}^1$ , qui est continue car  $h(1) - h(0) \in \mathbb{Z}$ . Alors pour tout  $x \in \mathbb{T}^1$ , nous avons  $p \cdot h(x) = h \circ f(x) \pmod{\mathbb{Z}}$ , donc  $\phi_p \circ \bar{h} = \bar{h} \circ f$ . D'où  $\bar{h}$  est une semi-conjugaison entre  $(\mathbb{T}^1, \phi_p)$  et  $(\mathbb{T}^1, f)$ .

L'application  $h$  est croissante. Si elle n'est pas strictement croissante, alors elle est constante sur un intervalle non réduit à un point  $[a, b]$ . La relation de semi-conjugaison dit alors que  $h$  est constante sur l'intervalle  $\tilde{f}([a, b])$ , et par récurrence sur l'intervalle  $\tilde{f}^n([a, b])$ . Comme  $\tilde{f}$  est strictement dilatante, l'application  $h$  est constante, ce qui contredit le fait que  $h(1) = 1 \neq 0 = h(0)$ . Par conséquent,  $\bar{h}$  est une application continue bijective du cercle dans lui-même, donc est un homéomorphisme par compacité.

On montre de manière analogue que si  $f : \mathbb{T}^1 \rightarrow \mathbb{T}^1$  est une application strictement dilatante de degré  $p \leq -2$ , alors  $f$  est topologiquement conjuguée à  $\phi_p$ .

## 9 Dynamique sur les espaces homogènes

Dans toute cette partie, nous fixons un entier  $N \in \mathbb{N} - \{0, 1\}$ . Nous avons étudié dans la partie 4 divers exemples de systèmes dynamiques sur le tore  $\mathbb{T}^N$ , qui est le quotient du groupe topologique  $G = \mathbb{R}^N$  par le sous-groupe discret  $\Gamma = \mathbb{Z}^N$ , en particulier pour l'action par translations de  $\mathbb{R}^N$  sur  $\mathbb{T}^N$ . Dans ce chapitre, nous étendons cette étude à des groupes topologiques non abéliens comme  $G = \mathrm{SL}_N(\mathbb{R})$ .<sup>163</sup>

L'exemple le plus simple de tel système dynamique sur  $\Gamma \backslash G$  s'interprète géométriquement comme le flot géodésique sur un quotient du plan hyperbolique, voir la partie 9.5, ainsi que le joli livre [Dal].

Le quotient  $\Gamma \backslash G$  n'est en général plus un groupe, mais un espace homogène sous l'action de  $G$ . Les séries de Fourier, qui jouaient un rôle si important pour le tore  $\mathbb{T}^N$  (voir la partie 4), seront remplacées dans cette partie 9 par les représentations unitaires de  $G$ , voir la sous-partie 9.2.

En effet, si  $G = \mathbb{R}^N$ ,  $\Gamma = \mathbb{Z}^N$  alors la famille des fonctions trigonométriques  $(e_m)_{m \in \mathbb{Z}^N}$  est une base hilbertienne de l'espace de Hilbert  $\mathcal{H} = \mathbb{L}^2(\mathbb{T}^N, d\theta)$ , sur lequel le groupe  $\mathbb{R}^N$  agit par  $(t, f) \mapsto (\pi(t)f : x \pmod{\mathbb{T}^N} \mapsto f(x + t \pmod{\mathbb{T}^N}))$ . Cette action est une représentation unitaire de  $G = \mathbb{R}^N$  sur  $\mathcal{H}$  (voir ci-dessous pour leur définition). La démonstration de l'ergodicité des translations par un élément  $t$  de  $\mathbb{R}^N$  totalement irrationnel sur  $\mathbb{T}^N$  (voir la proposition 4.3) passait par l'étude des coefficients matriciels (voir ci-dessous pour leur définition)  $\langle \pi(nt)e_m, e_{m'} \rangle$ , et en particulier leur annulation pour  $n$  assez grand. C'est ce résultat d'annulation qui formera la trame de tout ce chapitre (voir la proposition 9.9).

### 9.1 Le groupe unimodulaire $\mathrm{SL}_N(\mathbb{R})$

Notons  $G$  le groupe topologique localement compact, métrisable, séparable

$$G = \mathrm{SL}_N(\mathbb{R}) = \{g \in \mathrm{GL}_N(\mathbb{R}) : \det g = 1\}.$$

Il existe sur  $G$  une (en fait de très nombreuses, voir par exemple [Pau1, Coro. 1.3]) distance  $d_G$  invariante à gauche, c'est-à-dire telle que pour tous les  $g, x, y$  dans  $G$ , nous ayons

$$d_G(gx, gy) = d_G(x, y).$$

<sup>163</sup>. Cette étude se généralise à tous les groupes de Lie réels  $G$  connexes, semi-simples, de centre fini, voir par exemple [BekM].

Par exemple, en notant indifféremment  $e = \text{id} = I_N$  la matrice identité de taille  $N$  et  $\|\cdot\|$  une norme matricielle<sup>164</sup> quelconque sur  $\mathcal{M}_N(\mathbb{R})$ , nous pouvons prendre la distance<sup>165</sup>

$$d_G(x, y) = \log(1 + \|x^{-1}y - \text{id}\|) + \log(1 + \|y^{-1}x - \text{id}\|). \quad (69)$$

**Lemme 9.1.** (1) *Il existe une mesure borélienne positive non nulle  $\nu_G$  sur  $G$ , finie sur les compacts, invariante par les translations à gauche<sup>166</sup>, unique à un scalaire multiplicatif près.*

(2) *La mesure  $\mu$  est aussi invariante par les translations à droite<sup>167</sup>.*

**Remarque.** L'assertion (1) est vraie pour tout groupe topologique localement compact  $G$  (en renforçant "finie sur les compacts" par "régulière" si  $G$  n'est pas supposé métrisable séparable, voir [Coh]). La mesure  $\nu_G$  est appelée une *mesure de Haar* (à gauche) de  $G$ . Lorsque l'assertion (2) est vérifiée, nous dirons que  $G$  est *unimodulaire*. Par exemple un groupe topologique compact est unimodulaire. Voir par exemple [Wei] pour des démonstrations et informations supplémentaires sur les mesures de Haar.

**Démonstration.** Il est facile de construire explicitement  $\nu_G$  pour  $G = \text{SL}_N(\mathbb{R})$ . Notons  $\lambda$  la mesure de Lebesgue sur  $\mathcal{M}_N(\mathbb{R}) = \mathbb{R}^{N^2}$ , et posons, pour tout borélien  $A$  de  $G$ ,

$$\nu_{\text{SL}_N(\mathbb{R})}(A) = \lambda\left(\left\{x \in \mathcal{M}_N(\mathbb{R}) : 1 \leq \det x \leq 2 \text{ et } \frac{1}{(\det x)^{\frac{1}{N}}} x \in A\right\}\right).$$

Nous avons bien  $(\lambda_g)_*\nu_G = (\rho_g)_*\nu_G = \nu_G$ , car les endomorphismes linéaires de  $\mathcal{M}_N(\mathbb{R})$  définis par  $x \mapsto gx$  et  $x \mapsto xg^{-1}$  sont de déterminant 1, pour tout  $g$  dans  $G$ .<sup>168</sup>

Pour montrer l'unicité, soient  $\nu$  et  $\nu'$  deux mesures positives boréliennes non nulles, finies sur les compacts, invariantes par les translations à gauche sur  $G$ .

Notons que le support d'une mesure de Haar de  $G$  est égal à  $G$ . En effet, tout ouvert non vide  $U$  de  $G$  contient un translaté à gauche de la boule ouverte  $B_G(e, \epsilon)$  pour un  $\epsilon > 0$  assez petit. Comme  $G$  admet une partie  $P$  dénombrable dense et par invariance à gauche de la distance, nous avons  $G = \bigcup_{g \in P} gB_G(e, \epsilon)$ . Puisque  $\nu$  est non nulle et invariante, nous avons  $\nu(B_G(e, \epsilon)) \neq 0$ , donc  $\nu(U) \neq 0$ . Donc toute fonction continue à support compact, qui est positive et non identiquement nulle, est d'intégrale finie non nulle pour toute mesure de Haar.

Une mesure de Haar sur  $G$  est  $\sigma$ -finie car si  $V$  est un voisinage compact de  $e$ , alors  $G = \bigcup_{g \in P} gV$  par l'argument ci-dessus, et la mesure de Haar de  $V$  est finie.

164. c'est-à-dire une norme sur  $\mathcal{M}_N(\mathbb{R})$  vérifiant  $\|xy\| \leq \|x\| \|y\|$  pour tous les  $x, y \in \mathcal{M}_N(\mathbb{R})$ , comme la norme d'opérateur des matrices vues comme endomorphismes linéaires de l'espace euclidien standard  $\mathbb{R}^N$

165. L'inégalité triangulaire découle par exemple des inégalités

$$\begin{aligned} 1 + \|x^{-1}z - e\| &= 1 + \|x^{-1}y - e\| + \|y^{-1}z - e\| + (\|x^{-1}z - e\| - \|x^{-1}y - e\| - \|y^{-1}z - e\|) \\ &\leq 1 + \|x^{-1}y - e\| + \|y^{-1}z - e\| + \|(x^{-1}z - e) - (x^{-1}y - e) - (y^{-1}z - e)\| \\ &= 1 + \|x^{-1}y - e\| + \|y^{-1}z - e\| + \|(x^{-1}y - e)(y^{-1}z - e)\| \leq (1 + \|x^{-1}y - e\|)(1 + \|y^{-1}z - e\|). \end{aligned}$$

166. c'est-à-dire  $(\lambda_g)_*\nu_G = \nu_G$  pour tout  $g \in G$ , où  $\lambda_g : x \mapsto gx$  est la translation à gauche par  $g$

167. c'est-à-dire  $(\rho_g)_*\nu_G = \nu_G$  pour tout  $g \in G$ , où  $\rho_g : x \mapsto xg^{-1}$  est la translation à droite par  $g$

168. En effet, en considérant une matrice de  $\mathcal{M}_N(\mathbb{R})$  comme le  $N$ -uplet de ses vecteurs colonnes, et par un calcul de déterminant diagonal par blocs, pour tout  $g$  dans  $G$ , le déterminant de l'endomorphisme de  $(\mathbb{R}^N)^N$  défini par  $(X_1, \dots, X_N) \mapsto (gX_1, \dots, gX_N)$  est  $(\det g)^n$ . De plus,  $xg = {}^t({}^t g {}^t x)$  pour tout  $x \in \mathcal{M}_N(\mathbb{R})$  et le déterminant de l'endomorphisme linéaire involutif  $x \mapsto {}^t x$  de  $\mathcal{M}_N(\mathbb{R})$  vaut  $(-1)^{\frac{n(n-1)}{2}}$ .

Fixons  $g \in C_c^0(G)$  positive non identiquement nulle <sup>169</sup>. Pour tout  $f \in C_c^0(G)$ , notons

$$h : (x, y) \mapsto \frac{f(x)g(yx)}{\int_G g(tx) d\nu'(t)},$$

qui est une application de  $G \times G$  dans  $\mathbb{C}$  bien définie (le dénominateur ne s'annule pas par ce qui précède), à support compact dans  $G \times G$ , et continue (le dénominateur est continu en  $x$  par l'uniforme continuité de  $g$ ).

Par le théorème de Fubini (puisque les mesures sont  $\sigma$ -finies) et par l'invariance par translation à gauche de  $\nu$  par  $y^{-1}$  et de  $\nu'$  par  $x$ , nous avons

$$\begin{aligned} \int_G \int_G h(x, y) d\nu'(y) d\nu(x) &= \int_G \int_G h(x, y) d\nu(x) d\nu'(y) \\ &= \int_G \int_G h(y^{-1}x, y) d\nu(x) d\nu'(y) = \int_G \int_G h(y^{-1}x, y) d\nu'(y) d\nu(x) \\ &= \int_G \int_G h(y^{-1}, xy) d\nu'(y) d\nu(x). \end{aligned}$$

Donc

$$\begin{aligned} \int_G f(x) d\nu(x) &= \int_G f(x) \frac{\int_G g(yx) d\nu'(y)}{\int_G g(tx) d\nu'(t)} d\nu(x) \\ &= \int_G \int_G h(x, y) d\nu'(y) d\nu(x) = \int_G \int_G h(y^{-1}, xy) d\nu'(y) d\nu(x) \\ &= \int_G \int_G \frac{f(y^{-1})g(x)}{\int_G g(ty^{-1}) d\nu'(t)} d\nu'(y) d\nu(x) \\ &= \left( \int_G g(x) d\nu(x) \right) \left( \int_G \frac{f(y^{-1})}{\int_G g(ty^{-1}) d\nu'(t)} d\nu'(y) \right). \end{aligned}$$

Puisque l'intégrale de droite de la ligne ci-dessous ne dépend pas de la mesure  $\nu$ , nous avons donc

$$\frac{\int_G f(x) d\nu(x)}{\int_G g(x) d\nu(x)} = \frac{\int_G f(x) d\nu'(x)}{\int_G g(x) d\nu'(x)}.$$

En notant  $c = \frac{\int_G g(x) d\nu'(x)}{\int_G g(x) d\nu(x)} > 0$  qui ne dépend pas de  $f$ , nous avons par conséquent  $\int_G f(x) d\nu(x) = \int_G f(x) d(c\nu')(x)$ . Puisque deux mesures donnant même intégrale à toutes les fonctions continues à support compact coïncident, nous avons donc  $\nu = c\nu'$ .  $\square$

Nous fixons dans la suite une mesure de Haar  $\nu_G$  sur  $G$ . Soit  $\Gamma$  un sous-groupe discret de  $G$ . Il existe donc  $\epsilon_0 > 0$  tel que  $B_G(e, 2\epsilon_0) \cap \Gamma = \{e\}$ . Par définition, le quotient  $X = \Gamma \backslash G$  est l'ensemble des classes à gauche  $x = \Gamma g$  des éléments  $g \in G$ . Nous le munissons de la distance (qui induit la topologie quotient)

$$d(\Gamma g, \Gamma g') = \inf_{\gamma \in \Gamma} d_G(g, \gamma g'),$$

et nous notons  $\pi$  la projection canonique de  $G$  dans  $X = \Gamma \backslash G$  définie par  $g \mapsto \Gamma g$ .

169. par exemple la restriction à  $G$  d'une fonction dans  $C_c^0(\mathbb{R}^{N^2})$ , positive et valant 1 en  $I_N$



Nous notons encore  $\rho$  l'action de  $G$  sur  $X$  par translations à droite sur les classes à gauche modulo  $\Gamma$  : pour tout  $g$  dans  $G$ , nous avons, lorsque  $\tilde{x}$  parcourt  $G$ ,

$$\rho_g : x = \Gamma \tilde{x} \mapsto xg^{-1} = \Gamma \tilde{x} g^{-1} .$$

**Lemme 9.2.** (1) Pour tout  $g$  dans  $G$ , l'application  $\rho$  induit un homéomorphisme de  $B_G(g, \epsilon_0)$  sur  $B(\Gamma g, \epsilon_0)$ . De plus, l'image réciproque  $\rho^{-1}(B(\Gamma g, \epsilon_0))$  est la réunion disjointe des boules  $B_G(\gamma g, \epsilon_0)$  pour  $\gamma$  dans  $\Gamma$ .

(2) Il existe une mesure positive borélienne non nulle  $\mu = \mu_X$  sur  $X$ , finie sur les compacts et invariante par translations à droite<sup>170</sup>, unique à scalaire multiplicatif près.

**Remarque.** L'assertion (2) est encore vraie pour tous les groupes topologiques localement compacts unimodulaire  $G$  (en remplaçant "finie sur les compacts" par "régulière" si  $G$  n'est pas supposé métrisable séparable) et leurs sous-groupes discrets  $\Gamma$ . La mesure  $\mu$  est appelée une *mesure de Haar* (à droite) de l'espace homogène  $X = \Gamma \backslash G$ . Voir par exemple [Wei] pour des démonstrations et généralisations (en particulier lorsque  $\Gamma$  est remplacé par un sous-groupe fermé de  $G$ ). Attention, la mesure  $\mu$  n'est en général pas la mesure image de  $\nu_G$  par  $\rho$ . Par contre, la restriction de  $\mu$  à toute boule  $B(\Gamma g, \frac{\epsilon_0}{2})$  est la mesure image par  $\rho$  de la restriction d'un multiple scalaire strictement positif de  $\nu_G$  à  $B_G(g, \frac{\epsilon_0}{2})$ .

**Démonstration.** L'assertion (1) se démontre comme pour le tore  $\mathbb{T}^N$ , et implique que  $\rho$  est un revêtement (voir par exemple [Pau4]).

(2) Nous définissons d'abord  $\mu$  sur les boules  $B(\Gamma g, \frac{\epsilon_0}{2})$  : pour tout borélien  $A$  contenu dans  $B(\Gamma g, \frac{\epsilon_0}{2})$ , nous écrivons  $A = \rho(\tilde{A})$  avec  $\tilde{A}$  un borélien contenu dans  $B(g, \frac{\epsilon_0}{2})$  et posons  $\mu(A) = \nu_G(\tilde{A})$ . Cela ne dépend pas des choix, car  $\nu_G$  est invariante à gauche.

Définissons alors  $\mu$  sur  $X$  : pour tout borélien  $A$ , nous écrivons  $A = \bigsqcup_{i \in \mathbb{N}} A_i$  avec les  $A_i$  qui sont des boréliens deux à deux disjoints, de diamètre au plus  $\frac{\epsilon_0}{2}$ , et nous posons  $\mu(A) = \sum_{i \in \mathbb{N}} \mu(A_i)$ . Cela ne dépend pas non plus des choix. Par construction, et comme  $G$  est unimodulaire, pour tout  $g$  dans  $G$ , nous avons  $(\rho_g)_* \mu(A) = \mu((\rho_g)^{-1}(A)) = \mu(Ag) = \mu(A)$ .  $\square$

Un sous-groupe discret  $\Gamma$  de  $G$  est appelé un *réseau* si  $\mu(\Gamma \backslash G) < \infty$ , ce qui ne dépend pas du choix de  $\mu$ . Nous normaliserons alors la mesure de Haar  $\mu$  de l'espace homogène  $\Gamma \backslash G$  pour qu'elle soit de probabilité. Un sous-groupe discret  $\Gamma$  de  $G$  est dit *cocompact* (ou *uniforme*) si le quotient  $\Gamma \backslash G$  est compact ; c'est en particulier un réseau.

**Exemples :** (1) Le sous-groupe  $\mathrm{SL}_N(\mathbb{Z})$  est un réseau non uniforme de  $\mathrm{SL}_N(\mathbb{R})$ . Voir la proposition 9.16 ci-dessous pour  $N = 2$ .

(2) Le groupe topologique  $\mathrm{SL}_N(\mathbb{R})$  contient aussi des réseaux cocompacts.

Nous ne montrerons pas ces assertions, voir par exemple [Bor, BorHC]. Ces exemples sont construits de façon arithmétique et sont le point de départ d'une interaction riche entre systèmes dynamiques et arithmétique, voir par exemple [Zim, Mar1, Mar2].

170. c'est-à-dire telle que  $(\rho_g)_* \mu = \mu$  pour tout  $g \in G$

## 9.2 Représentations unitaires

Soit  $\Gamma$  un réseau de  $G$ . L'action par translation à droite d'un élément  $g$  de  $G$  sur  $X = \Gamma \backslash G$  est un système dynamique topologique, qui préserve la mesure de Haar  $\mu$  de  $X$  (qui est finie). Pour étudier les propriétés ergodiques de ce système dynamique, nous aurons besoin de la définition suivante. Nous renvoyons par exemple à [BeHV] pour des généralités sur cette partie.

Si  $\mathcal{H}$  est un espace de Hilbert (complexe),<sup>171</sup> notons  $U(\mathcal{H})$  le *groupe unitaire* de  $\mathcal{H}$ , c'est-à-dire le groupe des bijections linéaires continues<sup>172</sup>  $u : \mathcal{H} \rightarrow \mathcal{H}$

- qui préservent le produit scalaire de  $\mathcal{H}$  :

$$\forall x, y \in \mathcal{H}, \quad \langle u(x), u(y) \rangle = \langle x, y \rangle ,$$

- ou, de manière équivalente par les formules de polarisation, qui préservent la norme hilbertienne de  $\mathcal{H}$  :

$$\forall x \in \mathcal{H}, \quad \|u(x)\| = \|x\| ,$$

- ou, de manière équivalente, qui sont isométriques pour la distance hilbertienne  $d$  définie par  $(x, y) \mapsto \|y - x\|$  :

$$\forall x, y \in \mathcal{H}, \quad d(u(x), u(y)) = d(x, y) ,$$

- ou, de manière équivalente, dont l'inverse est égal à leur adjoint : nous avons  $uu^* = u^*u = 1$  où  $u^*$  est l'adjoint<sup>173</sup> de  $u$ .

Une *représentation unitaire* de  $G$  dans un espace de Hilbert  $\mathcal{H}$  est un morphisme de groupes  $\pi : G \rightarrow U(\mathcal{H})$  tel que, pour tout  $v$  dans  $\mathcal{H}$ , l'application  $G \rightarrow \mathcal{H}$  définie par  $g \mapsto \pi(g)v$  soit continue. En particulier, l'application de  $G \times \mathcal{H}$  dans  $\mathcal{H}$  définie par  $(g, v) \mapsto \pi(g)v$  est une action du groupe  $G$  sur l'ensemble  $\mathcal{H}$ . Nous noterons

$$\mathcal{H}^G = \{v \in \mathcal{H} : \forall g \in G, \pi(g)v = v\} .$$

le sous-espace vectoriel (qui est fermé) des vecteurs fixes de  $\mathcal{H}$  par cette action.

**Exemple :** La *représentation quasi-régulière* (droite) de  $G$  sur  $X = \Gamma \backslash G$  est la représentation unitaire  $\pi : G \rightarrow U(\mathbb{L}^2(X, \mu))$  de  $G$  dans l'espace de Hilbert  $\mathcal{H} = \mathbb{L}^2(X, \mu)$ , où, pour tous les  $g \in G$  et  $f \in \mathcal{H}$ , nous avons

$$\pi(g)f : x \mapsto f(xg) .$$

Le sous-espace vectoriel

$$\mathbb{L}_0^2(X, \mu) = \{f \in \mathbb{L}^2(X, \mu) : \int_X f d\mu = 0\}$$

est un sous-espace de Hilbert de  $\mathcal{H}$ , qui est invariant par  $G$ . Nous avons

$$\mathbb{L}_0^2(X, \mu)^G = \{0\} . \tag{70}$$

171. Son produit scalaire sera noté  $(x, y) \mapsto \langle x, y \rangle$ . Il est supposé linéaire à gauche et anti-linéaire à droite.

172. leur inverse est alors automatiquement continu par le théorème de Banach, mais cela découle aussi de la propriété suivante d'isométrie

173. Rappelons que l'adjoint  $u^*$  de  $u$  est l'unique application linéaire continue de  $\mathcal{H}$  dans  $\mathcal{H}$  telle que  $\langle u(x), y \rangle = \langle x, u^*(y) \rangle$  pour tous les  $x, y \in \mathcal{H}$ .

**Démonstration.** Montrons que  $\pi$  est en effet une représentation unitaire. Comme  $\mu$  est invariante par l'action par translations à droite de  $G$ , nous avons

$$\forall f, f' \in \mathbb{L}^2(X, \mu), \forall g \in G, \quad \int_X f(xg) \overline{f'(xg)} d\mu(x) = \int_X f(x) \overline{f'(x)} d\mu(x),$$

donc l'application linéaire  $\pi(g)$  appartient bien à  $U(\mathcal{H})$ . Il est immédiat que  $\pi$  est un morphisme de groupes.<sup>174</sup> Montrons la continuité, pour tout  $f \in \mathcal{H}$  fixé, de l'application  $g \mapsto \pi(g)f$ . C'est vrai si  $f$  appartient à l'espace vectoriel  $C_c^0(X)$  des fonctions continues à support compact sur  $X$ , par continuité uniforme. Nous concluons par la densité de  $C_c^0(X)$  dans  $\mathcal{H}$  de la manière suivante. Soient  $(g_n)_{n \in \mathbb{N}}$  une suite dans  $G$  qui converge vers un élément  $g \in G$  et  $\epsilon > 0$ . Il existe  $f_0$  dans  $C_c^0(X)$  tel que  $\|f - f_0\|_{\mathbb{L}^2} \leq \epsilon$ . Alors

$$\|\pi(g_n)f - \pi(g)f\|_{\mathbb{L}^2} \leq 2\|f - f_0\|_{\mathbb{L}^2} + \|\pi(g_n)f_0 - \pi(g)f_0\|_{\mathbb{L}^2} \leq 3\epsilon$$

pour  $n$  assez grand.

Comme la mesure  $\mu$  est finie, nous avons

$$\left| \int f d\mu \right| \leq \|f\|_1 \leq \|\mu\| \|f\|_2$$

par l'inégalité de Cauchy-Schwarz. Donc  $\mathbb{L}^2(X, \mu)$  est contenu dans  $\mathbb{L}^1(X, \mu)$ , et  $\mathbb{L}_0^2(X, \mu)$  est bien défini. Il est immédiat que le sous-espace  $\mathbb{L}_0^2(X, \mu)$  est fermé dans  $\mathbb{L}^2(X, \mu)$  par la continuité de la forme linéaire  $f \mapsto \int f d\mu$ , et qu'il est invariant par  $G$ .

Soit  $f \in \mathbb{L}_0^2(X, \mu)^G$ . Pour tout  $g$  dans  $G$ , nous avons  $f(xg) = f(x)$  pour  $\mu$ -presque tout  $x$ . Donc, par le théorème de Fubini, pour  $\mu$ -presque tout  $x$ , nous avons  $f(xg) = f(x)$  pour  $\nu_G$ -presque tout  $g$ . Fixons un tel point  $x$ . Lorsque  $g$  décrit le complémentaire d'un ensemble  $\nu_G$ -négligeable, les points  $xg$  décrivent le complémentaire d'un ensemble  $\mu$ -négligeable. Donc  $f$  est constante  $\mu$ -presque tout. Comme  $f$  est d'intégrale nulle, elle est nulle  $\mu$ -presque tout.  $\square$

Le résultat suivant découle de ce qui précède.

**Corollaire 9.3.** *Si  $\mu$  est une mesure de Haar sur l'espace homogène  $X = \Gamma \backslash G$ , alors  $\mathbb{L}^2(X, \mu)^G$  est réduit aux fonctions  $\mu$ -presque partout constantes.*  $\square$

Soit  $\pi : G \rightarrow U(\mathcal{H})$  une représentation unitaire de  $G$ . Pour tout  $v$  dans  $\mathcal{H}$ , notons

$$G_v = \{g \in G : \pi(g)v = v\}$$

le *fixateur* de  $v$ . Pour  $v, v'$  dans  $\mathcal{H}$ , notons  $c_{v,v'}$  l'application de  $G$  dans  $\mathbb{C}$  définie par

$$c_{v,v'}(g) = \langle \pi(g)v, v' \rangle,$$

appelée un *coefficient matriciel* de  $\pi$ . C'est une fonction continue sur  $G$ , uniformément bornée par  $\|v\| \|v'\|$  par l'inégalité de Cauchy-Schwarz. Le coefficient matriciel  $c_{v,v'}$  a comme propriété remarquable d'être invariant à droite par le fixateur  $G_v$  de  $v$  et invariant à gauche par le fixateur  $G_{v'}$  de  $v'$  :

$$\forall g, h, h' \in G, \quad \text{si } \pi(h)v = v \text{ et } \pi(h')v' = v', \quad \text{alors } c_{v,v'}(h'gh) = c_{v,v'}(g). \quad (71)$$

174. En effet, pour tous les  $g, g' \in G$ ,  $f \in \mathbb{L}^2(X, \mu)$  et  $x \in X$ , nous avons  $(\pi(e)f)(x) = f(x)$  et  $(\pi(gg')f)(x) = f(xgg') = (\pi(g')f)(xg) = ((\pi(g)\pi(g')f)(x))$ .

### 9.3 Ergodicité des translations sur les espaces homogènes

Dans cette partie, nous étudions les propriétés d'ergodicité de l'action par translation à droite de certains éléments de  $G = \mathrm{SL}_2(\mathbb{R})$  sur les espaces homogènes  $\Gamma \backslash G$  pour  $\Gamma$  un réseau de  $G$ . Cette étude sera généralisée dans la partie suivante.

Pour  $t, s \in \mathbb{R}$ , notons  $a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$ ,  $u_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}$ , et  $u_s^- = \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix}$ , qui sont des éléments de  $\mathrm{SL}_2(\mathbb{R})$ . Notons  $\mathrm{id} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ .

**Proposition 9.4.** *Soient  $\Gamma$  un réseau de  $G = \mathrm{SL}_2(\mathbb{R})$  et  $t, s \neq 0$ . Alors les translations à droite par  $a_t$ ,  $u_s$  et  $u_s^-$  sur  $X = \Gamma \backslash G$  sont ergodiques pour la mesure de Haar  $\mu$  de l'espace homogène  $X$ .*

**Démonstration.** D'après l'assertion (4) de la proposition 2.1, ce résultat est une conséquence de la proposition suivante, appliquée à  $\mathcal{H} = \mathbb{L}^2(X, \mu)$  et  $\pi$  la représentation quasi-régulière de  $G$  sur  $\mathcal{H}$ , en utilisant le corollaire 9.3.  $\square$

**Proposition 9.5.** *Soient  $(\mathcal{H}, \pi)$  une représentation unitaire de  $G$  et  $t, s \neq 0$ . Alors tout vecteur  $v$  de  $\mathcal{H}$  qui est invariant par  $a_t$  (respectivement invariant par  $u_s$  ou invariant par  $u_s^-$ ) est invariant par  $G$ .*

Cette proposition repose sur les lemmes suivants.

**Lemme 9.6. (Phénomène de Mautner)** *Soit  $v \in \mathcal{H}$ .*

- a) *Le fixateur  $G_v$  de  $v$  est un sous-groupe fermé de  $G$ .*
- b) *Nous avons  $g \in G_v$  si et seulement si  $c_{v,v}(g) = \|v\|^2$ .*
- c) *Soit  $g$  dans  $G$  tel qu'il existe des suites  $(g_n)_{n \in \mathbb{N}}$  dans  $G$  et  $(s_n)_{n \in \mathbb{N}}, (s'_n)_{n \in \mathbb{N}}$  dans  $G_v$  telles que*

$$\lim_{n \rightarrow \infty} g_n = g \quad \text{et} \quad \lim_{n \rightarrow \infty} s_n g_n s'_n = \mathrm{id} .$$

*Alors  $g$  appartient à  $G_v$ .*

**Démonstration.** L'assertion a) est claire par la propriété de continuité des représentations unitaires.

Montrons l'assertion b). Nous avons l'égalité

$$\|\pi(g)v - v\|^2 = 2(\|v\|^2 - \mathrm{Re}(c_{v,v}(g))) .$$

Si  $c_{v,v}(g) = \|v\|^2$ , alors en particulier  $c_{v,v}(g)$  est réel, et l'égalité ci-dessus montre que  $g \in G_v$ . Réciproquement, si  $g \in G_v$ , alors l'égalité ci-dessus montre que  $\mathrm{Re}(c_{v,v}(g)) = \|v\|^2$ . Or par l'inégalité de Cauchy-Schartz, nous avons

$$\sqrt{\mathrm{Re}(c_{v,v}(g))^2 + \mathrm{Im}(c_{v,v}(g))^2} = |c_{v,v}(g)| \leq \|\pi(g)v\| \|v\| = \|v\|^2 .$$

Donc  $\mathrm{Im}(c_{v,v}(g)) = 0$  et  $c_{v,v}(g) = \|v\|^2$ .

Pour montrer l'assertion c), remarquons que  $c_{v,v}(g_n) = c_{v,v}(s_n g_n s'_n)$  par la formule (71). Par continuité, nous en déduisons que  $c_{v,v}(g) = c_{v,v}(\mathrm{id}) = \|v\|^2$ . Il résulte de l'assertion b) que  $g$  appartient à  $G_v$ .  $\square$

Le phénomène de Mautner permet de montrer le lemme suivant.

**Lemme 9.7.** Soit  $v \in \mathcal{H}$ .

- (1) S'il existe  $t \neq 0$  tel que  $v$  est invariant par  $a_t$ , alors, pour tout  $s \in \mathbb{R}$ , le vecteur  $v$  est invariant par  $u_s$  et par  $u_s^-$ .
- (2) S'il existe  $s \neq 0$  tel que  $v$  est invariant par  $u_s$  (respectivement invariant par  $u_s^-$ ), alors pour tout  $t \in \mathbb{R}$ , le vecteur  $v$  est invariant par  $a_t$ .

**Démonstration.** (1) Nous pouvons supposer  $t > 0$ . Pour montrer que  $v$  est invariant par  $u_s$  pour tout  $s \in \mathbb{R}$ , appliquons le lemme 9.6 c) avec  $g_n = g = u_s$ ,  $s_n = a_t^{-n}$  et  $s'_n = a_t^n$ . Nous avons juste à vérifier que

$$s_n g_n s'_n = \begin{pmatrix} 1 & e^{-nt} s \\ 0 & 1 \end{pmatrix} \xrightarrow{n \rightarrow \infty} \text{id},$$

ce qui est clair. Nous montrons de même que  $v$  est invariant par  $u_s^-$  pour tout  $s \in \mathbb{R}$ .

(2) Supposons que  $v$  est invariant par  $u_s$ . Pour montrer que  $v$  est invariant par  $a_t$  pour tout  $t \in \mathbb{R}$ , il suffit par densité et par le lemme 9.6 a) de le prouver lorsque  $t$  est de la forme  $t = 2 \ln \frac{p}{q}$  avec  $\frac{p}{q}$  un rationnel non nul. Nous appliquons alors le lemme 9.6 c) avec

$$g = a_t = \begin{pmatrix} \frac{p}{q} & 0 \\ 0 & \frac{q}{p} \end{pmatrix}, \quad g_n = \begin{pmatrix} \frac{p}{q} & 0 \\ \frac{p-1}{snp} & \frac{q}{p} \end{pmatrix},$$

$$s_n = (u_s)^{-np} = \begin{pmatrix} 1 & -snp \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad s'_n = (u_s)^{nq} = \begin{pmatrix} 1 & snq \\ 0 & 1 \end{pmatrix}.$$

Comme  $\lim_{n \rightarrow \infty} g_n = g$ , nous avons seulement à vérifier que

$$s_n g_n s'_n = \begin{pmatrix} 1 & 0 \\ \frac{p-1}{snp} & 1 \end{pmatrix} \xrightarrow{n \rightarrow \infty} \text{id},$$

ce qui est clair. Nous procédons de la même façon lorsque  $v$  est invariant par  $u_s^-$ .  $\square$

**Démonstration de la proposition 9.5.** Soit  $v$  un vecteur de  $\mathcal{H}$  invariant par  $a_t$  pour un  $t \neq 0$  ou par  $u_s$  ou  $u_s^-$  pour un  $s \neq 0$ . Il résulte du lemme 9.7 que le groupe  $G_v$  fixant  $v$  contient tous les éléments de la forme  $a_{t'}$ ,  $u_{s'}$  ou  $u_{s'}^-$  pour  $t', s' \in \mathbb{R}$ . Comme ces éléments engendrent <sup>175</sup>  $G$ , nous avons  $G_v = G$ .  $\square$

## 9.4 Mélange des translations sur les espaces homogènes

Nous renvoyons par exemple à [BekM, HT] pour le contenu de cette sous-partie. L'étude précédente pour  $G = \text{SL}_2(\mathbb{R})$  va nous permettre de montrer le résultat plus général suivant. Comme déjà dit, nous munissons l'espace vectoriel réel de dimension fini  $\mathcal{M}_N(\mathbb{R})$  d'une norme matricielle quelconque.

---

175. Nous avons  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{c}{a} & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \frac{1}{a} \end{pmatrix} \begin{pmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{pmatrix}$  si  $a > 0$ . Si  $a < 0$ , nous nous ramenons au cas précédent en multipliant  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  par  $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^4$ . Si  $a = 0$ , nous nous ramenons au cas  $a > 0$  en multipliant  $\begin{pmatrix} 0 & b \\ c & d \end{pmatrix}$  par  $\begin{pmatrix} 0 & \mp 1 \\ \pm 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^{\pm 2}$ . Voir plus généralement le lemme 9.14.

**Théorème 9.8.** Soient  $\Gamma$  un réseau de  $G = \mathrm{SL}_N(\mathbb{R})$  et  $g$  un élément de  $G$  tel que  $\lim_{n \rightarrow \infty} \|g^n\| = \infty$ . Alors la translation à droite par  $g$  sur l'espace topologique  $X = \Gamma \backslash G$  est mélangeante pour la mesure de Haar  $\mu$  de  $X$ .

Une démonstration analogue montre que si  $(g^t)_{t \in \mathbb{R}}$  est un sous-groupe à un paramètre de  $G$  tel que  $\lim_{t \rightarrow +\infty} \|g^t\| = \infty$ , alors le système dynamique mesuré à temps continu  $(\Gamma \backslash G, \mu, (g^t)_{t \in \mathbb{R}})$  est mélangeant. Comme un système dynamique mesuré inversible  $(Y, \mu, \phi)$  est mélangeant si et seulement si son inverse  $(Y, \mu, \phi^{-1})$  l'est, le théorème 9.8 est valable pour  $g \in G$  tel que  $\lim_{n \rightarrow \infty} d_G(e, g^n) = +\infty$ .

Nous allons utiliser le théorème suivant de Howe<sup>176</sup>-Moore<sup>176</sup> de décroissance des coefficients des représentations unitaires de  $G$ . Le théorème de Howe-Moore, ainsi que le théorème 9.8, sont valables lorsque  $G$  est n'importe quel groupe de Lie réel connexe, semi-simple, de centre fini. Voir par exemple [BekM, Chap. III]).

**Proposition 9.9. (Théorème de Howe-Moore)** Soient  $G = \mathrm{SL}_N(\mathbb{R})$ ,  $\pi$  une représentation unitaire de  $G$  dans un espace de Hilbert  $\mathcal{H}$ , et  $v, w$  des vecteurs de  $\mathcal{H}$ . Si  $\mathcal{H}^G = \{0\}$ , alors

$$\lim_{\|g\| \rightarrow \infty} \langle \pi(g)v, w \rangle = 0.$$

**Démonstration du théorème 9.8.** Appliquons le théorème de Howe-Moore 9.9 avec  $\mathcal{H} = \mathbb{L}_0^2(\Gamma \backslash G, \mu)$  et  $\pi$  la restriction à  $\mathcal{H}$  de la représentation quasi-régulière de  $G$ . L'hypothèse de la proposition 9.9 est vérifiée par la formule (70). Le théorème 9.8 découle alors de la proposition 3.1 (5) car pour tous les  $v, w \in \mathcal{H}$ , le  $n$ -ème coefficient de corrélation  $c_n(v, w)$  est un coefficient matriciel de  $\pi$  :

$$c_n(v, w) = \int_X (v \circ (\rho_g)^n) w \, d\mu = \int_X v(xg^{-n}) w(x) \, d\mu(x) = \langle \pi(g^{-n})v, \bar{w} \rangle. \quad \square$$

Pour montrer le théorème de Howe-Moore 9.9, nous aurons besoin de deux résultats classiques. Notons  $A$  le sous-groupe de Cartan<sup>176</sup> de  $\mathrm{SL}_N(\mathbb{R})$ , constitué des matrices diagonales de déterminant 1 à coefficients diagonaux strictement positifs. Notons  $A^+$  la partie de  $A$ , définie par

$$A^+ = \left\{ \begin{pmatrix} t_1 & & 0 \\ & \ddots & \\ 0 & & t_N \end{pmatrix} : t_1 \geq \dots \geq t_N > 0 \text{ et } \prod_{i=1}^N t_i = 1 \right\}.$$

Notons  $K$  le sous-groupe compact maximal

$$K = \mathrm{SO}(N) = \{x \in \mathrm{GL}_N(\mathbb{R}) : x^{-1} = {}^t x\}$$

de  $\mathrm{SL}_N(\mathbb{R})$ .



Roger Howe  
176. 1945–



Calvin Moore  
1936–



Elie Cartan  
1869–1951



Shrikrishna Dani

**Lemme 9.10. (Décomposition de Cartan)** *Nous avons  $G = KA^+K$ .*

Autrement dit, tout élément  $g$  de  $G$  peut s'écrire  $g = k_1 a k_2$  avec  $k_1, k_2$  dans  $K$  et  $a$  dans  $A^+$ . Notons que  $a$  est unique (ses valeurs propres  $t_1, \dots, t_N$  dans cet ordre sont appelées les *valeurs singulières* de  $g$ , voir par exemple [Pau7, §1.10]. Par contre, les éléments  $k_1$  et  $k_2$  ne sont pas forcément uniques.

**Démonstration.** (Voir par exemple [MT].) Soit  $g \in G$ . La matrice  ${}^t g g$  est symétrique, définie positive. Il existe donc une matrice symétrique définie positive  $p$  telle que  ${}^t g g = p^2$ . En posant  $k' = g p^{-1}$ , nous avons  ${}^t k' k' = e$ , donc  $k' \in K$  et  $g = k' p$ . Nous pouvons diagonaliser  $p$  dans une base orthonormée en ordonnant ses valeurs propres de manière décroissante, d'où  $p = k a k^{-1}$  avec  $k$  dans  $K$  et  $a$  dans  $A^+$ . Nous avons alors  $g = k' k a k^{-1}$ . C'est ce que nous voulions.  $\square$

Rappelons qu'une suite  $(v_n)_{n \in \mathbb{N}}$  dans un espace de Hilbert  $\mathcal{H}$  converge faiblement vers un vecteur  $v$  dans  $\mathcal{H}$  si, pour tout  $w$  dans  $\mathcal{H}$ , nous avons

$$\lim_{n \rightarrow \infty} \langle v_n, w \rangle = \langle v, w \rangle .$$

Nous noterons alors  $v_n \rightharpoonup v$ . Le produit scalaire de  $\mathcal{H}$  étant non dégénéré, ce vecteur  $v$  est alors unique. Rappelons<sup>177</sup> qu'une application linéaire continue  $u : \mathcal{H} \rightarrow \mathcal{H}$  est faiblement continue : si  $v_n \rightharpoonup v$ , alors  $u(v_n) \rightharpoonup u(v)$ .

**Lemme 9.11. (Compacité faible des boules unités fermées des espaces de Hilbert)** *Toute suite bornée  $(v_n)_{n \in \mathbb{N}}$  dans un espace de Hilbert admet une sous-suite  $(v_{n_k})_{k \in \mathbb{N}}$  qui converge faiblement.*<sup>178</sup>

**Démonstration.** Quitte à remplacer  $\mathcal{H}$  par l'adhérence du sous-espace vectoriel engendré par  $\{v_n : n \in \mathbb{N}\}$ , nous pouvons supposer que  $\mathcal{H}$  est séparable. Soit  $(w_i)_{i \in \mathbb{N}}$  une partie dénombrable dense dans  $\mathcal{H}$ . Par un procédé d'extraction diagonale, nous pouvons supposer que pour tout  $i$ , la suite bornée  $(\langle v_n, w_i \rangle)_{n \in \mathbb{N}}$  converge. Nous en déduisons, par approximation, que pour tout  $w$  dans  $\mathcal{H}$ , la suite  $(\langle v_n, w \rangle)_{n \in \mathbb{N}}$  est de Cauchy, donc converge vers un réel  $f(w)$ . La majoration  $|f(w)| \leq \|w\| \max_{n \in \mathbb{N}} \|v_n\|$  (obtenue par l'inégalité de Cauchy-Schwarz) montre que  $f$  est une forme semi-linéaire continue. Donc par le théorème de dualité de Riesz<sup>179</sup>, elle est de la forme  $f : w \mapsto \langle v, w \rangle$  pour un  $v \in \mathcal{H}$ .  $\square$

**Démonstration du théorème de Howe-Moore 9.9.** Supposons par l'absurde qu'il existe  $\epsilon > 0$ , des vecteurs  $v', w'$  dans  $\mathcal{H}$  et une suite  $(g_n)_{n \in \mathbb{N}}$  dans  $G$  tels que

$$\forall n \in \mathbb{N}, \quad |\langle \pi(g_n) v', w' \rangle| \geq \epsilon \quad \text{et} \quad \lim_{n \rightarrow \infty} \|g_n\| = \infty .$$

177. En effet, pour tout  $w \in \mathcal{H}$ , nous avons

$$\lim_{n \rightarrow \infty} \langle u(v_n), w \rangle = \lim_{n \rightarrow \infty} \langle v_n, u^*(w) \rangle = \langle v, u^*(w) \rangle = \langle u(v), w \rangle .$$

178. Le théorème de compacité faible des boules unités fermées des espaces de Hilbert dit plus précisément qu'un fermé borné d'un espace de Hilbert  $\mathcal{H}$  est compact pour la topologie faible sur  $\mathcal{H}$  (la plus petite topologie sur  $\mathcal{H}$  rendant continue les formes linéaires continues (pour la norme) sur  $\mathcal{H}$ ). Rappelons que la boule unité fermée de  $\mathcal{H}$  est métrisable pour la topologie faible si et seulement si  $\mathcal{H}$  est séparable. Voir par exemple [Pau3, §6.2].

179. Notons  $\overline{\mathcal{H}}$  l'espace vectoriel complexe conjugué de l'espace vectoriel complexe  $\mathcal{H}$  (c'est-à-dire dont la multiplication externe  $(\lambda, x) \mapsto \lambda x$  est remplacée par  $(\lambda, x) \mapsto \overline{\lambda} x$ ). Le théorème de dualité de Riesz dit que l'application de  $\mathcal{H}$  dans  $(\overline{\mathcal{H}})^*$  donnée par  $x \mapsto \{y \mapsto \langle x, y \rangle\}$  est un isomorphisme linéaire isométrique entre la norme de  $\mathcal{H}$  et la norme duale de la norme de  $\mathcal{H}$  (qui est aussi une norme sur  $\overline{\mathcal{H}}$ ), voir par exemple [Rud2] ou [Pau3, §6.2].

Par le lemme 9.11, quitte à extraire une sous-suite, nous pouvons supposer que la suite (de norme constante)  $(\pi(g_n)v')_{n \in \mathbb{N}}$  converge faiblement vers un vecteur  $v'_0$ . Un passage à la limite dans l'inégalité  $|\langle \pi(g_n)v', w' \rangle| \geq \epsilon$  assure que  $v'_0$  est non nul.

Par la décomposition de Cartan, écrivons  $g_n = k_n a_n k'_n$  avec  $k_n, k'_n$  dans  $K$  et  $a_n$  dans  $A^+$ . Comme  $K$  est métrisable compact, quitte à extraire, nous pouvons supposer que les limites  $k = \lim_{n \rightarrow \infty} k_n$  et  $k' = \lim_{n \rightarrow \infty} k'_n$  existent. Posons  $v = \pi(k')v'$  et  $v_0 = \pi(k^{-1})v'_0$ .

**Lemme 9.12.** *Avec ces notations, la suite  $(a_n)_{n \in \mathbb{N}}$  de  $A^+$  vérifie*

$$\lim_{n \rightarrow \infty} \|a_n\| = \infty$$

et la suite  $(\pi(a_n)v)_{n \in \mathbb{N}}$  converge faiblement vers le vecteur non nul  $v_0$ .

**Démonstration.** La première assertion est claire. Pour montrer la seconde, pour tout  $w$  dans  $\mathcal{H}$ , en posant  $w'' = \pi(k)w$ , puisque  $\pi$  est une représentation unitaire, nous avons

$$|\langle \pi(a_n)v - v_0, w \rangle| = |\langle \pi(ka_n k')v' - v'_0, w'' \rangle| \leq T_1 + T_2 + T_3$$

avec

$$T_1 = |\langle \pi(ka_n k')v' - \pi(ka_n k'_n)v', w'' \rangle| \leq \|\pi(k')v' - \pi(k'_n)v'\| \|w''\| \xrightarrow{n \rightarrow \infty} 0,$$

$$\begin{aligned} T_2 &= |\langle \pi(ka_n k'_n)v' - \pi(k_n a_n k'_n)v', w'' \rangle| = |\langle \pi(a_n k'_n)v', \pi(k^{-1})w'' - \pi(k_n^{-1})w'' \rangle| \\ &\leq \|v'\| \|\pi(k^{-1})w'' - \pi(k_n^{-1})w''\| \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

et

$$T_3 = |\langle \pi(g_n)v' - v'_0, w'' \rangle| \xrightarrow{n \rightarrow \infty} 0. \quad \square$$

Soient  $i, j \in \{1, \dots, N\}$ . Notons  $E_{ij}$  la matrice  $N \times N$  dont les coefficients sont tous nuls sauf le coefficient d'indice  $(i, j)$  qui vaut 1. Pour  $i \neq j$ , notons  $U_{ij}, A_{ij}, S_{ij}$  les sous-groupes suivants de  $G$  :

$$U_{ij} = \{\text{id} + sE_{ij} : s \in \mathbb{R}\}$$

$$A_{ij} = \{\text{id} + (t-1)E_{ii} + (t^{-1}-1)E_{jj} : t > 0\}$$

$$S_{ij} = \{\text{id} + (a-1)E_{ii} + bE_{ij} + cE_{ji} + (d-1)E_{jj} : ad - bc = 1\}$$

Le sous-groupe  $U_{ij}$  est abélien unipotent (isomorphe au groupe additif  $(\mathbb{R}, +)$  par l'isomorphisme  $s \mapsto \text{id} + sE_{ij}$ ). Le sous-groupe  $A_{ij}$  est abélien diagonal (isomorphe au groupe multiplicatif  $(\mathbb{R}_+^*, \times)$  par l'isomorphisme  $s \mapsto \text{id} + (t-1)E_{ii} + (t^{-1}-1)E_{jj}$ ). Le groupe  $S_{ij}$  est isomorphe au groupe  $\text{SL}_2(\mathbb{R})$  (par l'isomorphisme  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \text{id} + (a-1)E_{ii} + bE_{ij} + cE_{ji} + (d-1)E_{jj}$ ). Nous pourrions lui appliquer la proposition 9.5.

**Lemme 9.13.** *Avec ces notations, il existe  $\ell \in \{1, \dots, N-1\}$  tel que, pour tous les entiers  $i, j$  tels que  $1 \leq i \leq \ell < j \leq N$ , le vecteur  $v_0$  soit invariant par  $U_{ij}$ .*



**Démonstration.** La matrice  $a_n \in A^+$  de la décomposition de Cartan de  $g_n$  est diagonale avec des coefficients diagonaux  $(t_{i,n})_{1 \leq i \leq N}$  décroissants. Comme  $\lim_{n \rightarrow \infty} \|a_n\| = \infty$  et  $\det a_n = 1$ , nous pouvons trouver un entier  $\ell \in \{1, \dots, N\}$  tel que  $\lim_{n \rightarrow \infty} t_{\ell,n}/t_{\ell+1,n} = \infty$ . Rappelons que si  $X$  est une matrice diagonale inversible, de coefficients diagonaux  $x_1, \dots, x_N$ , et si  $Y = (y_{ij})_{1 \leq i, j \leq N}$ , alors la matrice  $X^{-1}YX$  est égale à  $(\frac{x_j}{x_i} y_{ij})_{1 \leq i, j \leq N}$ .

Soit  $u$  dans  $U_{ij}$  avec  $1 \leq i \leq \ell < j \leq N$ . Nous voulons montrer que  $\pi(u)v_0 = v_0$ . Le choix de  $\ell$  assure que nous avons  $\lim_{n \rightarrow \infty} a_n^{-1}ua_n = e$ . Par un calcul analogue à celui du lemme 9.12, nous avons alors (les limites étant prises au sens de la convergence faible sur  $\mathcal{H}$ , en rappelant que les applications linéaires continues sur  $\mathcal{H}$  sont faiblement continues)

$$\pi(u)v_0 = \lim_{n \rightarrow \infty} \pi(u)\pi(a_n)v = \lim_{n \rightarrow \infty} \pi(a_n)\pi(a_n^{-1}ua_n)v = \lim_{n \rightarrow \infty} \pi(a_n)v = v_0 .$$

C'est ce que nous voulions. □

**Lemme 9.14.** *Le groupe  $G = \mathrm{SL}_N(\mathbb{R})$  est engendré par les sous-groupes  $A_{ij}$  et  $U_{ij}$  pour  $1 \leq i, j \leq N$  et  $i \neq j$ .*

**Démonstration.** Pour  $g$  dans  $G$ , la matrice  $(\mathrm{id} + sE_{ij})g$  est obtenue à partir de  $g$  par « une opération élémentaire sur les lignes » qui consiste à ajouter à la  $i$ -ème ligne de  $g$  un multiple de sa  $j$ -ème ligne. De même, la matrice  $g(\mathrm{id} + sE_{ij})$  est obtenue à partir de  $g$  par « une opération élémentaire sur les colonnes ». Partant de n'importe quelle matrice  $g \in G$ , nous nous ramènon, par une succession d'opérations élémentaires sur les lignes et les colonnes à une matrice diagonale dans  $G$ , qui est produit d'éléments de  $A_{i,i+1}$  pour  $i = 1, \dots, N - 1$ , à un nombre pair de signes près, qui sont traités comme dans la note de bas de page 175. □

Terminons maintenant la démonstration du théorème de Howe-Moore. Pour résumer ce qui précède, en raisonnant par l'absurde, nous avons construit un entier  $\ell \leq N - 1$  et un vecteur non nul  $v_0$  de  $\mathcal{H}$ , qui est invariant par tous les groupes  $U_{ij}$  avec  $1 \leq i \leq \ell < j \leq N$ . D'après la proposition 9.5, ce vecteur  $v_0$  est invariant par  $S_{ij}$ . Il est en particulier invariant par  $A_{ij}$  et donc invariant par toutes les matrices  $A_{i'j'}$  avec  $1 \leq i', j' \leq N$  et  $i' \neq j'$  (car si  $i', j' \leq \ell$ , alors  $A_{i'j'} \subset A_{i'n}A_{j'n}$  et si  $i', j' > \ell$ , alors  $A_{i'j'} \subset A_{1i'}A_{1j'}$ ). De nouveau par la proposition 9.5, ce vecteur est invariant par  $U_{i'j'}$ . Le lemme 9.14 ci-dessus montre alors que le vecteur non nul  $v_0$  est invariant par  $G$ , contradiction. □

## 9.5 Flots géodésique et horocyclique sur les surfaces hyperboliques

Nous renvoyons par exemple à [Dal] pour le contenu de cette partie. Voir aussi le livre [Kat], en particulier son premier chapitre, et [BekM, Chap. IV]. Le but de cette partie est de construire des exemples remarquables de systèmes dynamiques différentiables à temps continu, de nature géométrique, et d'étudier leurs propriétés ergodiques.

Il existe une façon géométrique très simple de « visualiser » le groupe  $\mathrm{SL}_2(\mathbb{R})$ , ou plus précisément le groupe quotient<sup>180</sup>  $G = \mathrm{PSL}_2(\mathbb{R})$  de  $\mathrm{SL}_2(\mathbb{R})$  par son centre  $\{\pm \mathrm{id}\}$ . Un élément de  $G$  est, par définition, une matrice de déterminant 1 définie au signe près, que nous noterons  $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  avec  $a, b, c, d \in \mathbb{R}$  et  $ad - bc = 1$ .

180. Notons que la notation  $G$  dans cette partie ne désigne pas le même objet que les notations  $G$  des parties 9.1 à 9.4.

Nous munissons  $G$  d'une mesure de Haar (à gauche), de sorte que la projection canonique de  $\mathrm{SL}_2(\mathbb{R})$  sur  $\mathrm{PSL}_2(\mathbb{R})$  soit localement un isomorphisme mesuré pour les deux mesures de Haar. Le groupe  $G$  est aussi unimodulaire, ses mesures de Haar à gauche sont invariantes par les translations à droite. <sup>181</sup>

Pour tout sous-groupe discret  $\Gamma$  de  $G$ , il existe une mesure positive borélienne non nulle sur l'espace topologique quotient  $\Gamma \backslash G$ , invariante par translations à droite par  $G$ , unique modulo scalaire strictement positif, appelée la *mesure de Haar* de l'espace homogène  $\Gamma \backslash G$ . En fait, si  $\rho : \mathrm{SL}_2(\mathbb{R}) \rightarrow \mathrm{PSL}_2(\mathbb{R})$  est la projection canonique, alors  $\tilde{\Gamma} = \rho^{-1}(\Gamma)$  est un sous-groupe discret de  $\mathrm{SL}_2(\mathbb{R})$ , et  $\rho$  induit un homéomorphisme  $\tilde{\Gamma} \backslash \mathrm{SL}_2(\mathbb{R}) \rightarrow \Gamma \backslash G$ , qui préserve les mesures de Haar de ces espaces homogènes.

Un sous-groupe discret  $\Gamma$  de  $G$  est un *réseau* si la mesure de Haar de  $\Gamma \backslash G$  est finie. Il est dit *cocompact* (ou *uniforme*) si  $\Gamma \backslash G$  est compact. Nous normalisons alors la mesure de Haar de l'espace homogène  $\Gamma \backslash G$  pour être une mesure de probabilité. Notons que  $\Gamma$  est un réseau (respectivement réseau uniforme) de  $G$  si et seulement si  $\tilde{\Gamma}$  est un réseau (respectivement réseau uniforme) de  $\mathrm{SL}_2(\mathbb{R})$ .

Le groupe  $G$  agit sur le *demi-plan supérieur*

$$\mathbb{H} = \{z = x + iy \in \mathbb{C} : y = \mathrm{Im} z > 0\}$$

par la formule

$$\forall g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in G, \quad \forall z \in \mathbb{H}, \quad g \cdot z = \frac{az + b}{cz + d}.$$

En effet, cette formule ne dépend pas du signe de  $g$ ; un petit calcul montre que nous avons bien  $g \cdot (g' \cdot z) = (gg') \cdot z$ ; et le fait que  $G$  préserve  $\mathbb{H}$  découle du calcul

$$\mathrm{Im}(g \cdot z) = \frac{\mathrm{Im} z}{|cz + d|^2}. \quad (72)$$

Notons que  $z \mapsto g \cdot z$  est une application holomorphe sur l'ouvert  $\mathbb{H}$  de  $\mathbb{C}$ , dont la différentielle (en tant qu'application de classe  $C^1$ ) est

$$\xi \mapsto \frac{\xi}{(cz + d)^2}.$$

L'action de  $G$  s'étend continûment au compactifié d'Alexandrov  $\mathbb{R} \cup \{\infty\}$  de l'axe réel en posant, pour tous les  $g \in G$  et  $t \in \mathbb{R} \cup \{\infty\}$ ,

$$g \cdot t = \begin{cases} \frac{at+b}{ct+d} & \text{si } t \in \mathbb{R} \text{ et } t \neq -\frac{b}{c} \\ \infty & \text{si } t = \infty \text{ et } c = 0 \\ \frac{a}{c} & \text{si } t = \infty \text{ et } c \neq 0 \\ \infty & \text{si } t = -\frac{b}{c} \text{ et } c \neq 0. \end{cases}$$

Le *fibré tangent* à l'ouvert  $\mathbb{H}$  de  $\mathbb{C}$  est  $T\mathbb{H} = \mathbb{H} \times \mathbb{C}$ , réunion disjointe des espaces tangents  $T_z\mathbb{H} = \{z\} \times \mathbb{C}$  en tout point  $z \in \mathbb{H}$ . L'action tangente de  $G$  sur  $T\mathbb{H}$  est donnée par la formule

$$\forall g \in G, \quad \forall (z, \xi) \in T\mathbb{H}, \quad g \cdot (z, \xi) = \left( \frac{az + b}{cz + d}, \frac{\xi}{(cz + d)^2} \right). \quad (73)$$

---

181. Une manière de voir ceci est de dire que puisque le sous-groupe  $Z = \{\pm \mathrm{id}\}$  est discret et central, le groupe  $\mathrm{SL}_2(\mathbb{R})$  agit aussi par translations à gauche sur l'espace homogène  $Z \backslash \mathrm{SL}_2(\mathbb{R})$ , et une mesure de Haar de cet espace homogène est aussi invariante à gauche par  $\mathrm{SL}_2(\mathbb{R})$ .

Pour tout vecteur tangent  $v = (z, \xi) \in T\mathbb{H}$  où  $z = x + iy$ , notons

$$\|v\|_z = \frac{|\xi|}{\operatorname{Im} z} = \frac{|\xi|}{y},$$

et appelons cette quantité la *norme hyperbolique* de  $v$ .<sup>182</sup> Appelons *plan hyperbolique réel*, et notons  $\mathbb{H}_{\mathbb{R}}^2$ , l'ouvert  $\mathbb{H}$  muni de l'application norme hyperbolique de  $T\mathbb{H}$  dans  $\mathbb{R}$  définie par  $v = (z, \xi) \mapsto \|v\|_z$ . Il est élémentaire<sup>183</sup> de vérifier que l'action de  $G$  préserve la norme hyperbolique<sup>184</sup>, c'est-à-dire que pour tous les  $g \in G$  et  $v = (z, \xi) \in T\mathbb{H}$ , nous avons

$$\|g \cdot v\|_{g \cdot z} = \|v\|_z. \quad (74)$$

Le *volume hyperbolique* de  $\mathbb{H}_{\mathbb{R}}^2$  est la mesure  $\operatorname{vol}_{\mathbb{H}_{\mathbb{R}}^2}$  sur  $\mathbb{H}$  absolument continue par rapport à la mesure de Lebesgue de  $\mathbb{H}$ , définie par

$$d \operatorname{vol}_{\mathbb{H}_{\mathbb{R}}^2}(z) = \frac{dx \, dy}{y^2}.$$

Un calcul élémentaire<sup>185</sup> montre que l'action de  $G$  sur  $\mathbb{H}_{\mathbb{R}}^2$  préserve le volume hyperbolique. Notons

$$T^1\mathbb{H}_{\mathbb{R}}^2 = \{v = (z, \xi) \in \mathbb{H}_{\mathbb{R}}^2 \times \mathbb{C} : \|v\|_z = 1\},$$

que nous appelons le *fibré tangent unitaire* du plan hyperbolique réel  $\mathbb{H}_{\mathbb{R}}^2$ . Il est invariant par l'action du groupe  $G$ . L'application de  $\mathbb{R} \times ]0, +\infty[ \times \mathbb{R}/\mathbb{Z}$  dans le fibré tangent unitaire  $T^1\mathbb{H}_{\mathbb{R}}^2$  définie par  $(x, y, \theta) \mapsto (x + iy, y e^{2i\pi\theta})$  est un homéomorphisme, que nous utiliserons comme paramétrage local et global.

Si  $p : T^1\mathbb{H}_{\mathbb{R}}^2 \rightarrow \mathbb{H}_{\mathbb{R}}^2$  est la première projection  $v = (z, \xi) \mapsto z$ , alors la fibre de  $p$  au-dessus de  $z \in \mathbb{H}_{\mathbb{R}}^2$  est la *sphère tangente unité* en  $z$ , notée

$$T_z^1\mathbb{H}_{\mathbb{R}}^2 = \{v = (z, \xi) : |\xi| = \operatorname{Im} z\}.$$

La *mesure sphérique* sur  $T_z^1\mathbb{H}_{\mathbb{R}}^2$  est l'image de la mesure de Lebesgue du cercle  $\mathbb{R}/\mathbb{Z}$  par l'homéomorphisme  $\theta \mapsto (z, (\operatorname{Im} z) e^{2i\pi\theta})$ , notée  $\operatorname{vol}_{T_z^1\mathbb{H}_{\mathbb{R}}^2}(v)$ . Nous la considérerons aussi comme une mesure sur  $T^1\mathbb{H}_{\mathbb{R}}^2$ , à support dans  $T_z^1\mathbb{H}_{\mathbb{R}}^2$ .

La *mesure de Liouville* sur le fibré tangent unitaire  $T^1\mathbb{H}_{\mathbb{R}}^2$ , notée  $m_{\text{Liou}}$  est la mesure qui désintègre (au sens expliqué ci-dessous) par la projection  $p$  au-dessus du volume hyperbolique, avec mesures conditionnelles sur les fibres les mesures sphériques :

$$m_{\text{Liou}}(v) = \int_{z \in \mathbb{H}_{\mathbb{R}}^2} \operatorname{vol}_{T_z^1\mathbb{H}_{\mathbb{R}}^2}(v) \, d \operatorname{vol}_{\mathbb{H}_{\mathbb{R}}^2}(z).$$

182. Cette norme étant une norme euclidienne sur chaque espace tangent, variant de manière lisse, ceci définit une *métrique riemannienne* sur  $\mathbb{H}_{\mathbb{R}}^2$ , voir par exemple [GaHL, Pau5].

183. En effet, par les formules (73) et (72), nous avons

$$\|g \cdot v\|_{g \cdot z} = \frac{\left| \frac{\xi}{(cz+d)^2} \right|}{\operatorname{Im}(g \cdot z)} = \frac{|\xi|}{|cz+d|^2 \frac{\operatorname{Im} z}{|cz+d|^2}} = \frac{|\xi|}{\operatorname{Im} z} = \|v\|_z.$$

184. On dit que l'action de  $G$  est *isométrique* pour la métrique riemannienne susdite.

185. En effet,  $\frac{dx \wedge dy}{y^2} = \frac{i}{2} \frac{dz \wedge d\bar{z}}{(\operatorname{Im} z)^2}$  et par les formules (73) et (72),

$$\frac{d(g \cdot z) \wedge d(\overline{g \cdot z})}{(\operatorname{Im}(g \cdot z))^2} = \frac{\frac{dz}{(cz+d)^2} \wedge \frac{d\bar{z}}{(c\bar{z}+d)^2}}{\left(\frac{\operatorname{Im} z}{|cz+d|^2}\right)^2} = \frac{dz \wedge d\bar{z}}{(\operatorname{Im} z)^2}.$$

Ceci signifie que pour toute fonction continue à support compact  $f \in C_c^0(T^1\mathbb{H}_{\mathbb{R}}^2)$ , nous avons

$$\begin{aligned} \int_{T^1\mathbb{H}_{\mathbb{R}}^2} f(v) dm_{\text{Liou}}(v) &= \int_{z \in \mathbb{H}_{\mathbb{R}}^2} \left( \int_{v \in T_z^1\mathbb{H}_{\mathbb{R}}^2} f(v) d\text{vol}_{T_z^1\mathbb{H}_{\mathbb{R}}^2}(v) \right) d\text{vol}_{\mathbb{H}_{\mathbb{R}}^2}(z) \\ &= \int_{x \in \mathbb{R}} \int_{y \in ]0, +\infty[} \int_{\theta \in \mathbb{R}/\mathbb{Z}} f(x + iy, y e^{2i\pi\theta}) \frac{d\theta dx dy}{y^2}. \end{aligned} \quad (75)$$

Il est élémentaire de vérifier que la mesure de Liouville est une mesure positive borélienne non nulle sur  $T^1\mathbb{H}_{\mathbb{R}}^2$ , finie sur les compacts, invariante par l'action de  $G$ .

Le lemme suivant dit en particulier que  $G = \text{PSL}_2(\mathbb{R})$  agit simplement transitivement sur le fibré tangent unitaire du plan hyperbolique réel  $T^1\mathbb{H}_{\mathbb{R}}^2$ .

**Lemme 9.15.** Notons  $\Phi : G \rightarrow T^1\mathbb{H}_{\mathbb{R}}^2$  l'application définie par  $\Phi : g \mapsto g \cdot (i, i)$ .

- (1) L'application  $\Phi$  est un homéomorphisme.<sup>186</sup>
- (2) L'application  $\Phi$  est équivariante : pour tous les éléments  $g, g'$  dans  $G$ , nous avons  $\Phi(gg') = g \cdot \Phi(g')$ .
- (3) La mesure image de la mesure de Liouville sur  $T^1\mathbb{H}_{\mathbb{R}}^2$  par  $\Phi^{-1}$  est une mesure de Haar de  $G$ .

**Démonstration.** (1) et (2) L'action par homographies du groupe topologique  $G$  sur l'espace topologique  $\mathbb{H}_{\mathbb{R}}^2$  est continue. Elle est transitive, car par une translation horizontale  $z \mapsto z + t$  pour un certain  $t \in \mathbb{R}$  (correspondant à l'action de  $\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$ ), tout point de  $\mathbb{H}_{\mathbb{R}}^2$  s'envoie sur un point de l'axe vertical, et par une homothétie  $z \mapsto \lambda^2 z$  pour un certain  $\lambda > 0$  (correspondant à l'action de  $\begin{bmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{bmatrix}$ ), tout point de l'axe vertical s'envoie sur le point  $i$ .

Le stabilisateur du point  $i$  dans  $G$  est constitué des éléments  $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$  tels que  $a, b, c, d \in \mathbb{R}$ ,  $ad - bc = 1$  et  $\frac{ai+b}{ci+d} = i$ , donc avec  $a = d$ ,  $b = -c$  et  $a^2 + b^2 = 1$ . Il est donc égal à

$$\text{PSO}_2(\mathbb{R}) = \text{SO}_2(\mathbb{R})/\{\pm \text{id}\} = \left\{ \rho_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} : \theta \in \mathbb{R}/2\pi\mathbb{Z} \right\}.$$

Comme  $T_i^1\mathbb{H}_{\mathbb{R}}^2 = \{(i, \xi) : |\xi| = 1\}$  et puisque par la formule (73) nous avons

$$\rho_\theta \cdot (i, \xi) = \left( i, \frac{\xi}{((\sin \theta)i + \cos \theta)^2} \right) = (i, e^{-2i\theta} \xi),$$

le groupe  $\text{PSO}_2(\mathbb{R})$  agit simplement transitivement sur  $T_i^1\mathbb{H}_{\mathbb{R}}^2$ . Donc  $G$  agit simplement transitivement sur le fibré tangent unitaire  $T^1\mathbb{H}_{\mathbb{R}}^2$ .

Par conséquent, l'application orbitale  $\Phi : g \mapsto g \cdot (i, i)$  est une bijection continue équivariante. Il n'est pas difficile de montrer qu'elle est propre, donc est un homéomorphisme.

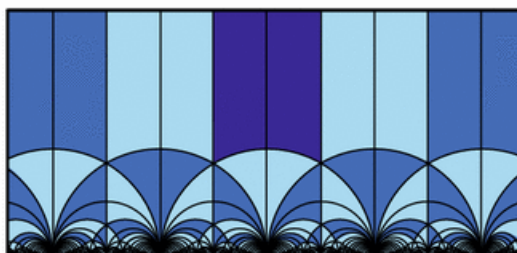
<sup>186.</sup> en fait un  $C^\infty$ -difféomorphisme quand  $G$  et  $T^1\mathbb{H}_{\mathbb{R}}^2$  sont munies des structures de variété différentielle quotient de la sous-variété différentielle  $\text{SL}_2(\mathbb{R})$  par  $\{\pm \text{id}\}$  et de sous-variété différentielle de  $\mathbb{H} \times \mathbb{C}$ , respectivement

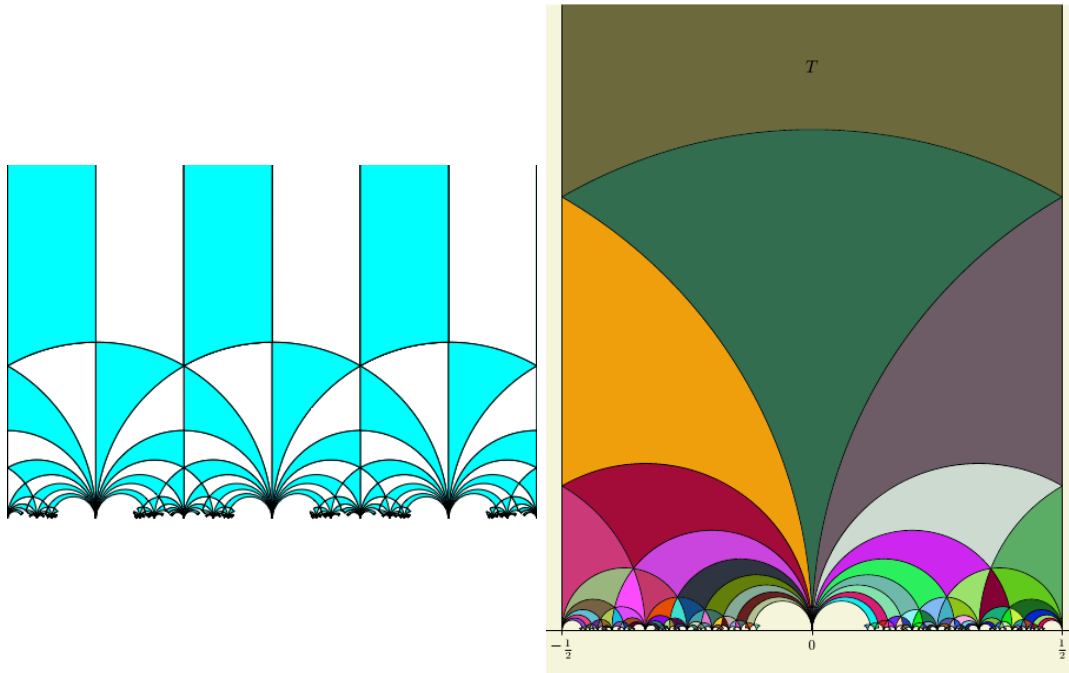
L'assertion (3) découle alors des assertions (1) et (2), car celles-ci montrent que la mesure image  $(\Phi^{-1})_* m_{\text{Liou}}$  est une mesure positive borélienne non nulle sur  $G$ , finie sur les compacts, invariante par l'action par translations à gauche de  $G$  sur lui-même.  $\square$

Soit  $\Gamma$  un sous-groupe discret de  $G$ . L'homéomorphisme équivariant  $\Phi$  induit par passage au quotient un homéomorphisme de  $\Gamma \backslash G$  dans  $\Gamma \backslash T^1 \mathbb{H}_{\mathbb{R}}^2$ , que nous notons encore  $\Phi$ . La mesure de Liouville sur  $T^1 \mathbb{H}_{\mathbb{R}}^2$  induit une mesure sur  $\Gamma \backslash T^1 \mathbb{H}_{\mathbb{R}}^2$ , de sorte que la projection canonique de  $T^1 \mathbb{H}_{\mathbb{R}}^2$  sur  $\Gamma \backslash T^1 \mathbb{H}_{\mathbb{R}}^2$  soit localement un isomorphisme mesuré, que l'on appelle encore la *mesure de Liouville* de  $\Gamma \backslash T^1 \mathbb{H}_{\mathbb{R}}^2$ . Sa mesure image par l'inverse de  $\Phi : \Gamma \backslash G \rightarrow \Gamma \backslash T^1 \mathbb{H}_{\mathbb{R}}^2$  est la mesure de Haar de l'espace homogène  $\Gamma \backslash G$  (voir le lemme 9.2 et la remarque qui le suit, ainsi que le début de la partie 9.5).

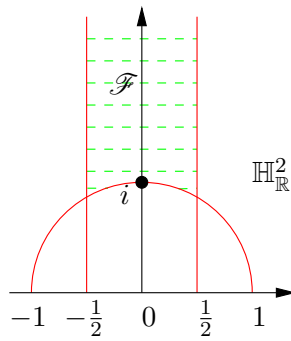
**Proposition 9.16.** *Le sous-groupe discret  $\Gamma = \text{PSL}_2(\mathbb{Z})$ , appelé le groupe modulaire, est un réseau de  $G = \text{PSL}_2(\mathbb{R})$ .*

Ceci montre aussi que  $\text{SL}_2(\mathbb{Z})$  est un réseau de  $\text{SL}_2(\mathbb{R})$ . Avant de fournir une démonstration, donnons quelques exemples de dessins de pavages du plan hyperbolique réel invariants par le groupe modulaire. Le premier est la bibliothèque de Richard Pink, les autres sont dus respectivement à J. Díaz-A. Verjovsky-F. Vlacci, P. McCreary-T. Murphy-C. Carter et P. Fradin.





**Démonstration.** Par l'assertion (3) du lemme 9.15, il suffit de montrer que la mesure de Liouville de  $\Gamma \backslash T^1 \mathbb{H}_{\mathbb{R}}^2$  est finie.



Par la méthode du pivot de Gauss, il est élémentaire de montrer que le groupe  $\mathrm{PSL}_2(\mathbb{Z})$  est engendré par  $S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$  et  $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ . Comme le groupe infini cyclique engendré par  $T$  agit sur  $\mathbb{H}_{\mathbb{R}}^2$  par les translations horizontales entières, et  $S$  par l'inversion  $z \mapsto -\frac{1}{z}$  par rapport au cercle unité, nous avons donc que

$$\mathcal{F} = \left\{ z \in \mathbb{H} : |z| \geq 1, \quad |\mathrm{Re} z| \leq \frac{1}{2} \right\}$$

est un *domaine (faiblement) fondamental*<sup>187</sup> pour l'action de  $\Gamma$  sur  $\mathbb{H}_{\mathbb{R}}^2$ . Voir l'exercice E.55.

187. c'est-à-dire que  $\bigcup_{\gamma \in \Gamma} \gamma \mathcal{F} = \mathbb{H}_{\mathbb{R}}^2$ . Nous avons aussi, en notant  $\overset{\circ}{\mathcal{F}} = \{z \in \mathbb{H} : |z| > 1, \quad |\mathrm{Re} z| < \frac{1}{2}\}$  l'intérieur de  $\mathcal{F}$ , les égalités  $\gamma \overset{\circ}{\mathcal{F}} \cap \gamma' \overset{\circ}{\mathcal{F}} = \emptyset$  si  $\gamma \neq \gamma'$  (voir l'exercice E.55), mais nous n'en aurons pas

Par désintégration de la mesure de Liouville sur  $T^1\mathbb{H}_{\mathbb{R}}^2$  (c'est-à-dire par la formule (75)), et puisque les mesures conditionnelles  $\text{vol}_{T_z^1\mathbb{H}_{\mathbb{R}}^2}$  sur les fibres  $T_z^1\mathbb{H}_{\mathbb{R}}^2 = p^{-1}(z)$  de  $p$  pour  $z \in \mathbb{H}_{\mathbb{R}}^2$  sont des mesures de probabilité (en tant que lois images de la mesure de Lebesgue sur  $\mathbb{R}/\mathbb{Z}$ ), nous avons <sup>188</sup>

$$m_{\text{Liou}}(\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2) \leq m_{\text{Liou}}(T^1\mathcal{F}) \leq \text{vol}_{\mathbb{H}_{\mathbb{R}}^2}(\mathcal{F}).$$

Or <sup>189</sup>

$$\text{vol}_{\mathbb{H}_{\mathbb{R}}^2}(\mathcal{F}) = \int_{\mathcal{F}} \frac{dx dy}{y^2} \leq \int_{y=\frac{1}{2}}^{\infty} \int_{x=-\frac{1}{2}}^{\frac{1}{2}} \frac{dx dy}{y^2} \leq 2.$$

Ceci montre le résultat.  $\square$

Nous définissons la *longueur (hyperbolique)*  $\ell(c)$  d'une courbe  $c : [a, b] \rightarrow \mathbb{H}_{\mathbb{R}}^2$ , où  $a \leq b$ , de classe  $C^1$  par morceaux <sup>190</sup>, dont nous noterons  $c'(t)$  le vecteur vitesse (la dérivée) en  $t \in [a, b]$  sauf pour un nombre fini de tels  $t$ , par la formule

$$\ell(c) = \int_a^b \|c'(t)\|_{c(t)} dt.$$

Par le théorème de changement de variable pour la mesure de Lebesgue, la longueur de  $c$  est invariante par changement de paramètre monotone à la source : pour toute application  $\varphi : [a', b'] \rightarrow [a, b]$  de classe  $C^1$  par morceaux et monotone, nous avons

$$\ell(c \circ \varphi) = \ell(c). \quad (76)$$

Par la formule (74), l'action au but par homographies de  $G$  sur les courbes  $c$  dans  $\mathbb{H}_{\mathbb{R}}^2$  de classe  $C^1$  par morceaux préserve leur longueur :

$$\forall g \in G, \quad \ell(g \cdot c) = \ell(c). \quad (77)$$

Une courbe  $c : I \rightarrow \mathbb{H}_{\mathbb{R}}^2$  de classe  $C^1$ , où  $I$  est un intervalle (par forcément borné) de  $\mathbb{R}$ , est dite *paramétrée par longueur d'arcs* (hyperbolique) si  $\|c'(t)\|_{c(t)} = 1$  pour tout  $t \in I$ , de sorte pour tous les  $s < t$  dans  $I$ , nous avons  $\ell(c|_{[s,t]}) = t - s$ .

Nous définissons la *distance (hyperbolique)*  $d(z, w)$  entre deux points  $z$  et  $w$  de  $\mathbb{H}_{\mathbb{R}}^2$  comme la borne inférieure des longueurs des courbes  $C^1$  par morceaux entre  $z$  et  $w$ . Il

besoin ici. Il est aussi possible de montrer (voir [Ser]) que

$$\mathcal{F}' = \left\{ z \in \mathcal{F} : \text{Re } z \neq \frac{1}{2} \text{ et } z \neq e^{i\theta} \text{ pour tout } \theta \in \left[ \frac{\pi}{3}, \frac{\pi}{2} \right] \right\}$$

est un *domaine fondamental strict*, c'est-à-dire contenant un et un seul point de toute orbite de  $\Gamma$  dans  $\mathbb{H}_{\mathbb{R}}^2$ .

188. En fait, ces deux inégalités sont des égalités. Elles découlent du fait que l'ensemble des points  $z$  de  $\mathbb{H}_{\mathbb{R}}^2$  dont le stabilisateur dans  $\Gamma$  est non trivial est de mesure nulle (cet ensemble est en fait discret, réunion des deux orbites de  $i$  et  $e^{\frac{2i\pi}{3}}$ , voir l'exercice E.55), du fait déjà mentionné que les images de l'intérieur du domaine fondamental  $\mathcal{F}$  par les éléments du groupe dénombrable  $\Gamma$  sont deux à deux disjointes, et parce que la frontière de  $\mathcal{F}$  est de mesure nulle.

189. Un calcul plus précis laissé au lecteur montre que  $\text{vol}_{\mathbb{H}_{\mathbb{R}}^2}(\mathcal{F}) = \frac{\pi}{3}$ .

190. c'est-à-dire continue et telle qu'il existe  $n \in \mathbb{N}$  et une subdivision finie  $a = t_0 \leq t_1 \leq \dots \leq t_n \leq t_{n+1} = b$  de l'intervalle  $[a, b]$  telle que la restriction de  $c$  à tout intervalle  $[t_i, t_{i+1}]$  pour  $i \in \llbracket 0, n \rrbracket$  soit de classe  $C^1$

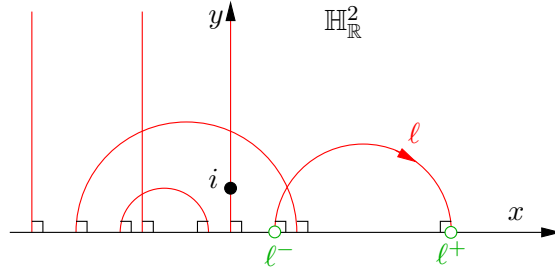
est facile de vérifier que  $d$  est bien une distance<sup>191</sup>, et par la formule (77) et passage aux bornes inférieures qu'elle est invariante par l'action de  $G$  sur  $\mathbb{H}_{\mathbb{R}}^2$ .

Appelons *géodésique (hyperbolique)* toute courbe continue  $c : I \rightarrow \mathbb{H}_{\mathbb{R}}^2$ , avec  $I$  un intervalle quelconque de  $\mathbb{R}$  (pas forcément borné), telle que pour tous les  $s < t$  dans  $I$ , nous avons

$$d(c(t), c(s)) = t - s.$$

Si  $I = [0, +\infty[$ , nous dirons que  $c$  est un *rayon géodésique* de  $\mathbb{H}_{\mathbb{R}}^2$  d'origine  $c(0)$ . Notons que l'image  $g \cdot c : s \mapsto g \cdot (c(s))$  d'une géodésique  $c : I \rightarrow \mathbb{H}_{\mathbb{R}}^2$  par l'action par homographie de tout élément  $g \in G$  est encore une géodésique.

**Lemme 9.17.** *Les géodésiques sont les arcs de demi-droites (ouvertes maximales) verticales et demi-cercles (ouverts maximaux) contenus dans  $\mathbb{H}_{\mathbb{R}}^2$  et orthogonaux à l'axe réel, paramétrés de manière  $C^\infty$  par longueur d'arc hyperbolique.*



Les deux *points à l'infini* d'une géodésique  $c$  de  $\mathbb{H}_{\mathbb{R}}^2$  définie sur  $\mathbb{R}$  sont les points d'intersection de l'axe réel avec le cercle contenant (l'image de)  $c$ , si  $c$  est un demi-cercle. Si  $c$  est une demi-droite verticale, ses *points à l'infini* sont le point  $\infty$  et le point d'intersection de l'axe réel avec la droite contenant (l'image de)  $c$ . On définit de manière similaire le *point à l'infini* d'un rayon géodésique de  $\mathbb{H}_{\mathbb{R}}^2$ . L'espace à l'infini de  $\mathbb{H}_{\mathbb{R}}^2$ , noté  $\partial_\infty \mathbb{H}_{\mathbb{R}}^2$ , est l'espace  $\mathbb{R} \cup \{\infty\}$  des points à l'infini des rayons géodésiques de  $\mathbb{H}_{\mathbb{R}}^2$ .

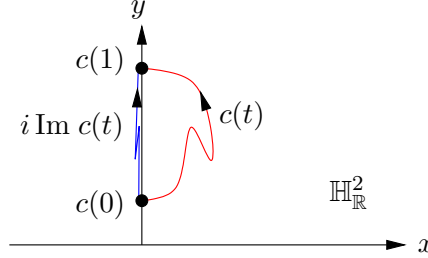
191. L'application  $d : \mathbb{H}_{\mathbb{R}}^2 \times \mathbb{H}_{\mathbb{R}}^2 \rightarrow \mathbb{R}$  est clairement bien définie par la connexité de l'ouvert  $\mathbb{H}$ , positive ou nulle, et nulle sur la diagonale. Si  $c : [a, a'] \rightarrow \mathbb{H}_{\mathbb{R}}^2$  est un chemin  $C^1$  par morceaux entre  $z$  et  $z'$  et  $\gamma : [a', a''] \rightarrow \mathbb{H}_{\mathbb{R}}^2$  est un chemin  $C^1$  par morceaux entre  $z'$  et  $z''$ , alors le chemin concaténé  $c \cdot \gamma : [a, a''] \rightarrow \mathbb{H}_{\mathbb{R}}^2$  valant  $c(t)$  en  $t \in [a, a']$  et  $\gamma(t)$  en  $t \in [a', a'']$  est encore  $C^1$  par morceaux (ceci est la raison pour la restriction à cette régularité), et de longueur  $\ell(c \cdot \gamma) = \ell(c) + \ell(\gamma)$  par l'additivité de l'intégrale. L'inégalité triangulaire de  $d$  s'en déduit par l'invariance de la longueur des courbes par translation à la source et par passage aux bornes inférieures.

Si  $c : [a, b] \rightarrow \mathbb{H}_{\mathbb{R}}^2$  est un chemin  $C^1$  par morceaux de  $z$  à  $w$ , nous avons déjà vu que l'application  $\gamma : [-a, -b] \rightarrow \mathbb{H}_{\mathbb{R}}^2$  définie par  $t \mapsto c(-t)$  est un chemin  $C^1$  par morceaux de  $w$  à  $z$  de même longueur que  $c$ . Donc  $d(w, z) \leq d(z, w)$  par passage aux bornes inférieures, avec égalité par argument de symétrie.

La fonction  $z \mapsto \frac{1}{\text{Im } z}$  est bornée et minorée par une constante strictement positive sur tout compact  $K$  de  $\mathbb{H}$ . Donc il existe  $c_K > 0$  telles les normes hyperboliques  $\|v\|_z = \frac{|\xi|}{\text{Im } z}$  et euclidiennes  $\|v\| = |\xi|$  des vecteurs  $v = (z, \xi)$  en tout point  $z \in K$  vérifient  $\frac{1}{c_K} \|v\| \leq \|v\|_z \leq c_K \|v\|$ . Donc pour tout chemin  $c$  de classe  $C^1$  par morceaux contenu dans  $K$ , les longueurs hyperbolique  $\ell(c) = \int_a^b \|c'(t)\|_{c(t)} dt$  et euclidienne  $\ell_{\text{euc}}(c) = \int_a^b \|c'(t)\| dt$  vérifient  $\frac{1}{c_K} \ell_{\text{euc}}(c) \leq \ell(c) \leq c_K \ell_{\text{euc}}(c)$ . Soient  $z$  et  $w$  deux points distincts de  $\mathbb{H}$ . Soit  $\epsilon > 0$  tel que la distance euclidienne entre  $z$  et  $w$  soit au moins  $2\epsilon$  et tel que la boule fermée euclidienne  $K$  de centre  $z$  et de rayon  $\epsilon$  soit contenue dans  $\mathbb{H}$ . Par le théorème des valeurs intermédiaires, tout chemin continu de  $z$  à  $w$  doit passer par le cercle euclidien  $C$  de centre  $z$  et de rayon  $\epsilon$ . La longueur euclidienne de tout chemin  $C^1$  par morceaux entre  $z$  et un point de  $C$  est au moins  $\epsilon$ , donc sa longueur hyperbolique est au moins  $\frac{\epsilon}{c_K}$ . Donc par passage à la borne inférieure, la distance hyperbolique entre  $z$  et  $w$  est au moins  $\frac{\epsilon}{c_K} > 0$ . Ceci montre la propriété de séparation de la distance hyperbolique de  $\mathbb{H}_{\mathbb{R}}^2$ .



**Démonstration.** Pour tout chemin  $c$  de classe  $C^1$  par morceaux entre deux points  $c(0)$  et  $c(1)$  de l'axe imaginaire, comme  $|c'(t)| \geq |(\operatorname{Im} c)'(t)|$ , nous avons  $\ell(c) \geq \ell(i \operatorname{Im} c)$ , avec inégalité stricte si la courbe  $c$  n'est pas contenue dans l'axe imaginaire.



De plus, tout chemin  $c : [a, b] \rightarrow \mathbb{H}_{\mathbb{R}}^2$  de classe  $C^1$  par morceaux sur l'axe imaginaire faisant un aller-retour non trivial<sup>192</sup> est de longueur (hyperbolique) strictement supérieure à celle de ce chemin privé de cet aller-retour. Donc la longueur de tout chemin  $C^1$  par morceaux entre deux points de l'axe imaginaire qui n'est pas monotone le long du segment entre ces deux points est strictement supérieure à celle de tout chemin  $C^1$  par morceaux et monotone entre ces deux points, qui ne dépend pas d'un tel chemin par la formule (76). Par conséquent, pour tout  $y > 1$ , la courbe  $\gamma : [1, y] \rightarrow \mathbb{H}_{\mathbb{R}}^2$  définie par  $s \mapsto s i$  étant  $C^1$  et monotone de  $i$  à  $iy$ , nous avons

$$d(i, yi) = \ell(\gamma) = \int_0^y \frac{ds}{s} = \log y .$$

De plus, la courbe  $c : [0, \log y] \rightarrow \mathbb{H}_{\mathbb{R}}^2$  définie par

$$t \mapsto e^t i \tag{78}$$

est alors l'unique (modulo translation à la source) géodésique  $C^1$  par morceaux joignant  $i$  à  $yi$ . Par approximation, nous pouvons facilement montrer que c'est aussi la seule géodésique (continue) joignant  $i$  à  $yi$  (et elle est donc de classe  $C^\infty$ ). Comme tout élément  $g$  de  $G$  est une isométrie hyperbolique, la courbe  $t \mapsto g \cdot (e^t i)$  est l'unique (modulo translation à la source) géodésique joignant  $g \cdot i$  à  $g \cdot (yi)$ .

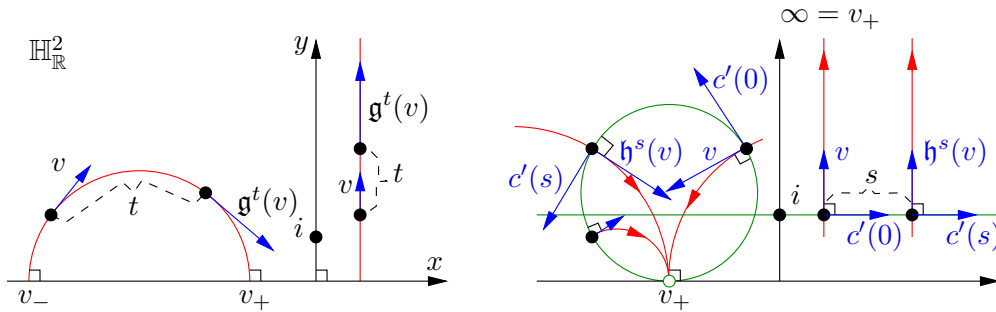
Il est bien connu (voir [Rud2]) que les homographies  $z \mapsto \frac{az+b}{cz+d}$  avec  $a, b, c, d$  dans  $\mathbb{C}$  et  $ad - bc = 1$  préservent la famille des cercles-droites, ainsi que la relation d'orthogonalité, qu'elles préservent le demi-plan supérieur si  $a, b, c, d$  sont réels, et qu'elles agissent de manière transitive sur la famille des cercles-droites. Le résultat en découle.  $\square$

Pour tout  $v = (z, \xi)$  dans  $\mathbb{T}^1 \mathbb{H}_{\mathbb{R}}^2$ , il existe donc une unique géodésique  $t \mapsto c(t)$ , définie sur  $\mathbb{R}$ , telle que  $v = (c(0), c'(0))$ . Notons alors, pour tout  $t \in \mathbb{R}$ ,

$$\mathbf{g}^t(v) = (c(t), c'(t)) ,$$

qui, si  $t \geq 0$ , est le vecteur tangent unitaire, tangent à la géodésique  $c$  au point à distance hyperbolique  $t$  de  $c(0)$  sur cette géodésique, voir le dessin de gauche ci-dessous.

<sup>192.</sup> c'est-à-dire tel qu'il existe  $t_0 < t_1 < t_2$  dans  $[a, b]$  tel que  $c(t_1)$  n'appartienne pas au segment entre  $c(t_0)$  et  $c(t_2)$



Appelons *horocycles*<sup>193</sup> les droites horizontales contenues dans  $\mathbb{H}_{\mathbb{R}}^2$ , dites *centrées* en  $\infty$ , ou les cercles contenus dans le demi-plan supérieur fermé, tangents à l'axe horizontal et privé de leur point sur cet axe, dits *centrés* en ce point (voir le dessin ci-dessus à droite). Le *point à l'infini* d'un tel horocycle est  $\infty$  dans le premier cas et le point de tangence dans le second cas.

Le groupe  $G$  des homographies  $z \mapsto \frac{az+b}{cz+d}$  avec  $a, b, c, d \in \mathbb{R}$  et  $ad - bc = 1$  agit transitivement sur l'ensemble des horocycles : les translations  $z \mapsto z + b$  permettent de ramener tout horocycle centré en un point de l'axe horizontal sur un horocycle centré en l'origine 0, l'homographie  $z \mapsto -\frac{1}{z}$  permet de transformer tout horocycle centré en l'origine sur un horocycle centré en  $\infty$ , et les homothéties  $z \mapsto \lambda^2 z$  avec  $\lambda > 0$  agissent transitivement sur les horocycles centrés en  $\infty$ . Les horocycles sont perpendiculaires aux géodésiques passant par leur point à l'infini, car c'est vrai pour les horocycles centrés à l'infini, et c'est vrai par transitivité pour les autres. L'application

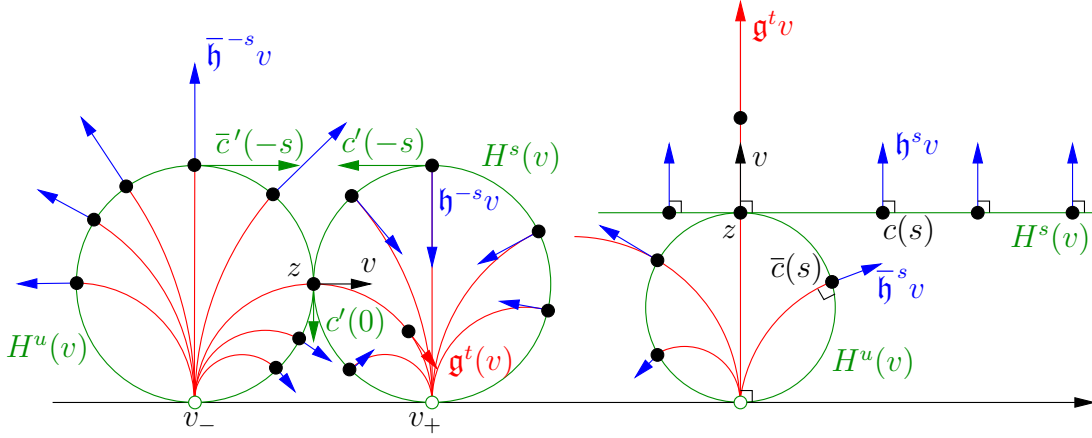
$$c : s \mapsto i + s \tag{79}$$

est l'unique paramétrage lisse par longueur d'arc hyperbolique de l'horocycle horizontal d'équation  $\text{Im } z = 1$  tel que  $(c(0), i c'(0)) = (i, i)$  : les vecteurs tangents  $(c(s) = i + s, c'(s) = 1)$  sont des vecteurs de norme hyperbolique 1.

Soit  $v = (z, \xi) \in T^1 \mathbb{H}_{\mathbb{R}}^2$  un vecteur tangent unitaire à  $\mathbb{H}_{\mathbb{R}}^2$ . Appelons *horocycle stable* de  $v$ , et notons  $H^s(v)$ , l'unique horocycle passant par l'origine  $z$  de  $v$  et centré au point en l'infini  $v_+$  du rayon géodésique défini par  $v$ . Appelons *horocycle instable* de  $v$ , et notons  $H^u(v)$ , l'unique horocycle passant par l'origine  $z$  de  $v$  et centré au point en l'infini  $v_-$  du rayon géodésique défini par  $(z, -\xi)$ . Voir le dessin ci-dessous, celui de gauche étant le cas où  $v_{\pm} \neq \infty$ . Notons que pour tous les  $g \in G$  et  $v \in T^1 \mathbb{H}_{\mathbb{R}}^2$ , nous avons

$$g \cdot H^s(v) = H^s(g \cdot v) \quad \text{et} \quad g \cdot H^u(v) = H^u(g \cdot v) .$$

193. Poincaré disait *horicycle*.



Flots horocycliques stables et instables

Pour tous les  $v \in T^1\mathbb{H}_{\mathbb{R}}^2$  et  $s \in \mathbb{R}$ , notons (voir le dessin ci-dessus)

$$\mathfrak{h}^s(v) = (c(s), i c'(s)) \in T^1\mathbb{H}_{\mathbb{R}}^2, \quad (80)$$

où  $s \mapsto c(s)$  est l'unique paramétrage lisse par longueur d'arc hyperbolique de l'horocycle stable  $H^s(v)$  de  $v$  tel que  $v = (c(0), i c'(0))$  soit le vecteur tangent unitaire positivement orthogonal au vecteur tangent  $(c(0), c'(0))$  de l'horocycle. De même notons

$$\bar{\mathfrak{h}}^s(v) = (\bar{c}(s), i \bar{c}'(s)) \in T^1\mathbb{H}_{\mathbb{R}}^2, \quad (81)$$

où  $s \mapsto \bar{c}(s)$  est l'unique paramétrage lisse par longueur d'arc hyperbolique de l'horocycle instable  $H^u(v)$  de  $v$  tel que  $v = (\bar{c}(0), i \bar{c}'(0))$ .

**Lemme 9.18.** *Les familles  $(\mathfrak{g}^t)_{t \in \mathbb{R}}$  et  $(\mathfrak{h}^s)_{s \in \mathbb{R}}$  sont des groupes à un paramètre de  $C^\infty$ -difféomorphismes de  $T^1\mathbb{H}_{\mathbb{R}}^2$ , appelés respectivement le flot géodésique et le flot horocyclique (stable) sur le plan hyperbolique réel  $\mathbb{H}_{\mathbb{R}}^2$ , qui commutent<sup>194</sup> avec l'action de  $G$  sur  $T^1\mathbb{H}_{\mathbb{R}}^2$ . Pour tous les  $g \in G$  et  $t, s \in \mathbb{R}$ , nous avons*

$$\mathfrak{g}^t(\Phi(g)) = \Phi(g a_t) \quad \text{et} \quad \mathfrak{h}^s(\Phi(g)) = \Phi(g u_s) \quad (82)$$

où  $a_t = \begin{bmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{bmatrix}$  et  $u_s = \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix}$ .

La mesure de Liouville sur  $T^1\mathbb{H}_{\mathbb{R}}^2$  est invariante par le flot géodésique et par le flot horocyclique.

**Démonstration.** La première assertion est immédiate. Pour la seconde, il suffit, par équivariance, de vérifier que nous avons

- l'égalité  $\mathfrak{g}^t(i, i) = \Phi(a_t)$ , ce qui est clair car l'unique géodésique définie sur  $\mathbb{R}$  de vecteur tangent  $(i, i)$  au temps  $t = 0$  est  $t \mapsto e^t i$  par la formule (78), donc les deux termes de cette égalité valent  $(e^t i, e^t i)$ ,

- l'égalité  $\mathfrak{h}^s(i, i) = \Phi(u_s)$ , ce qui est clair car l'unique paramétrage lisse par longueur d'arc hyperbolique de l'horocycle stable  $H^s(i, i) = \{z \in \mathbb{H}_{\mathbb{R}}^2 : \text{Im } z = 1\}$ , tel que  $(i, i)$  soit le vecteur tangent unitaire positivement orthogonal au vecteur vitesse au temps  $t = 0$  de

194. pour tous les  $t, s \in \mathbb{R}$ ,  $g \in G$  et  $v \in T^1\mathbb{H}_{\mathbb{R}}^2$ , nous avons  $\mathfrak{g}^t(g \cdot v) = g \cdot \mathfrak{g}^t(v)$  et  $\mathfrak{h}^s(g \cdot v) = g \cdot \mathfrak{h}^s(v)$

ce paramétrage, est  $s \mapsto i + s$  par la formule (79), donc les deux termes de cette égalité valent  $(i + s, i)$ .

La dernière assertion découle de la deuxième, du lemme 9.15 (3), et de l'invariance par translations à droite des mesures de Haar de  $G$ .  $\square$

Pour tout sous-groupe discret  $\Gamma$  de  $G$ , les flots  $(\mathfrak{g}^t)_{t \in \mathbb{R}}$  et  $(\mathfrak{h}^s)_{s \in \mathbb{R}}$  sur  $T^1\mathbb{H}_{\mathbb{R}}^2$  induisent par passage au quotient des flots sur  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ , appelés respectivement le *flot géodésique* sur  $\Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$  et le *flot horocyclique* sur  $\Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$ , et que nous notons encore  $(\mathfrak{g}^t)_{t \in \mathbb{R}}$  et  $(\mathfrak{h}^s)_{s \in \mathbb{R}}$ . Par sa construction et la dernière affirmation du lemme 9.18, la mesure de Liouville sur  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  est invariante par les flots géodésiques et horocycliques sur  $\Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$ .

En conclusion, pour tout réseau  $\Gamma$  de  $G$ , le flot sur l'espace homogène  $\Gamma \backslash G$  défini par la multiplication à droite par  $a_t$  pour  $t \in \mathbb{R}$  (respectivement  $u_s$  pour  $s \in \mathbb{R}$ ), étudié (modulo passage au quotient par le centre de  $\mathrm{SL}_2(\mathbb{R})$ ) dans la proposition 9.4, est  $C^\infty$ -conjugué au flot géodésique (respectivement horocyclique) sur le quotient  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ , et en particulier les diagrammes suivants commutent :

$$\begin{array}{ccc} \Gamma \backslash G & \xrightarrow{a_t} & \Gamma \backslash G & & \Gamma \backslash G & \xrightarrow{u_s} & \Gamma \backslash G \\ \Phi \downarrow & & \downarrow \Phi & & \Phi \downarrow & & \downarrow \Phi \\ \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2 & \xrightarrow{\mathfrak{g}^t} & \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2 & & \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2 & \xrightarrow{\mathfrak{h}^s} & \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2 \end{array} .$$

**Corollaire 9.19.** *Si  $\Gamma$  est un réseau de  $\mathrm{PSL}_2(\mathbb{R})$ , alors les flots géodésiques  $(\mathfrak{g}^t)_{t \in \mathbb{R}}$  et horocycliques  $(\mathfrak{h}^s)_{s \in \mathbb{R}}$  sur  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  sont mélangeants pour la mesure de Liouville de  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ .*

**Démonstration.** Notons  $\tilde{\Gamma}$  l'image réciproque de  $\Gamma$  par la projection canonique de  $\mathrm{SL}_2(\mathbb{R})$  sur  $\mathrm{PSL}_2(\mathbb{R})$ . Alors  $\tilde{\Gamma}$  est un réseau de  $\mathrm{SL}_2(\mathbb{R})$ , et la projection canonique induit un homéomorphisme  $\Theta$  de  $\tilde{\Gamma} \backslash \mathrm{SL}_2(\mathbb{R})$  sur  $\Gamma \backslash \mathrm{PSL}_2(\mathbb{R})$ , qui conjugue la translation à droite par  $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  à la translation à droite par  $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ , pour tout  $g \in \mathrm{SL}_2(\mathbb{R})$ . Par l'assertion (3) du lemme 9.15, la mesure image de la mesure de Liouville sur  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  par  $\Theta^{-1} \circ \Phi^{-1}$  est une mesure de Haar de l'espace homogène  $\tilde{\Gamma} \backslash \mathrm{SL}_2(\mathbb{R})$ . Par l'invariance de la propriété de mélange par une conjugaison de systèmes dynamiques mesurés, le résultat découle alors du théorème 9.8.  $\square$

Le résultat suivant découle du fait que mélangeant implique ergodique (voir la remarque (iii) suivant la proposition 3.1), du fait que la mesure de Liouville sur  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  est de support total, et de l'exercice E.12.

**Corollaire 9.20.** *Si  $\Gamma$  est un réseau de  $\mathrm{PSL}_2(\mathbb{R})$ , alors les flots géodésique  $(\mathfrak{g}^t)_{t \in \mathbb{R}}$  et horocyclique  $(\mathfrak{h}^s)_{s \in \mathbb{R}}$  sur  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  sont ergodiques pour la mesure de Liouville  $m_{\mathrm{Liou}}$  de  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ . L'orbite de  $m_{\mathrm{Liou}}$ -presque tout point de  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  par le flot géodésique ou par le flot horocyclique est dense.  $\square$*

**Exercice E.54.** *Soit  $\Gamma$  un sous-groupe discret de  $\mathrm{PSL}_2(\mathbb{R})$ .*

(1) *Montrer la propriété de commutation suivante des flots géodésique et horocyclique sur  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  : pour tous les  $t, s \in \mathbb{R}$ , nous avons*

$$\mathfrak{g}^t \circ \mathfrak{h}^s \circ \mathfrak{g}^{-t} = \mathfrak{h}^{s e^{-t}} . \quad (83)$$

(2) Montrer que la famille  $(\bar{h}^s)_{s \in \mathbb{R}}$  définie par la formule (81) est un groupe à un paramètre de  $C^\infty$ -difféomorphismes de  $T^1\mathbb{H}_{\mathbb{R}}^2$ . Il est appelé le flot horocyclique instable sur le plan hyperbolique réel  $\mathbb{H}_{\mathbb{R}}^2$ . Montrer qu'il commute avec l'action de  $G$  sur  $T^1\mathbb{H}_{\mathbb{R}}^2$ , qu'il préserve la mesure de Liouville sur  $T^1\mathbb{H}_{\mathbb{R}}^2$ , et qu'il vérifie, pour tous les  $g \in G$  et  $s, t \in \mathbb{R}$ ,

$$g^t \circ \bar{h}^s \circ g^{-t} = \bar{h}^{se^t} \quad \text{et} \quad \bar{h}^s(\Phi(g)) = \Phi(gu_s^-)$$

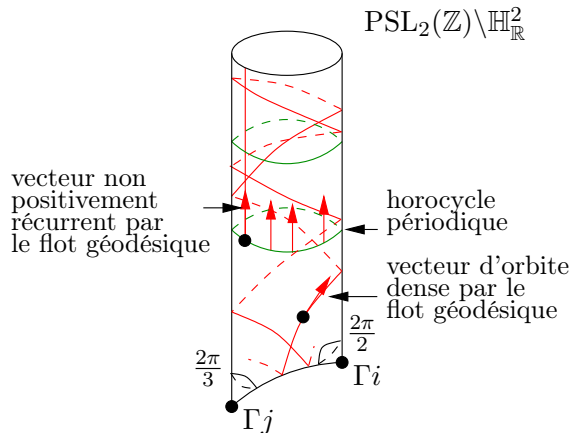
où  $u_s^- = \begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix}$ . Montrer que si  $\Gamma$  est un réseau de  $\mathrm{PSL}_2(\mathbb{R})$ , alors le flot horocyclique instable  $(\bar{h}^s)_{s \in \mathbb{R}}$  sur  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  défini par passage au quotient, est mélangeant pour la mesure de Liouville de  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ .

(3) Si  $\Gamma$  est un réseau de  $\mathrm{PSL}_2(\mathbb{R})$ , donner une autre démonstration de l'ergodicité du flot géodésique sur  $Y = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ , en considérant une fonction  $f \in \mathbb{L}^2(Y, m_{\mathrm{Liou}})$  invariante par le flot géodésique (c'est-à-dire telle que  $f \circ g^t = f$  dans  $\mathbb{L}^2(Y, m_{\mathrm{Liou}})$  pour tout  $t \in \mathbb{R}$ ), en montrant que  $\|f \circ h^s - f\|_{\mathbb{L}^2} = 0$  et  $\|f \circ \bar{h}^s - f\|_{\mathbb{L}^2} = 0$  pour tout  $s \in \mathbb{R}$  en utilisant la formule (83), et en montrant que  $f$  est invariante par l'action par translations à droite de  $G$  sur  $Y$  identifié à  $\Gamma \backslash G$  par  $\Phi^{-1}$ .

## 9.6 Exercices

**Exercice E.55.** Notons  $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$  le groupe modulaire,  $M = \Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$  la courbe modulaire<sup>195</sup> et  $T^1M = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ .<sup>196</sup> Notons  $\mathcal{F} = \{z \in \mathbb{C} : |z| \geq 1, |\mathrm{Re} z| \leq \frac{1}{2}\}$  le domaine (faiblement) fondamental usuel du groupe modulaire  $\Gamma$  sur  $\mathbb{H}_{\mathbb{R}}^2$ . Notons  $S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$  et

$$T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$



(1) Montrer que si deux points distincts  $z$  et  $z'$  dans  $\mathcal{F}$  sont dans la même orbite par  $\Gamma$ , alors ou bien  $\mathrm{Re} z = \pm \frac{1}{2}$  et  $z = z' \pm 1$ , ou bien  $|z| = 1$  et  $z' = -\frac{1}{z}$ . Montrer que le

195. Elle est souvent appelée la *surface modulaire* quand le demi-plan supérieur  $\mathbb{H}$  est considéré comme une sous-variété réelle de dimension 2, mais il vaut mieux le considérer comme une sous-variété complexe de dimension 1.

196. Nous renvoyons par exemple à [Paul] pour une description topologique de l'espace topologique quotient  $T^1M$ , homéomorphe au complémentaire du noeud de trèfle dans la sphère  $\mathbb{S}_3$ .

stabilisateur  $\Gamma_z = \{\gamma \in \Gamma : \gamma \cdot z = z\}$  d'un point  $z \in \mathcal{F}$  dans  $\Gamma$  est trivial (réduit à  $\{\text{id}\}$ ) sauf dans les trois cas suivants :

- $z = i$ , auquel cas  $\Gamma_z = \{\text{id}, S\}$ ,
- $z = \omega = e^{\frac{2i\pi}{3}}$ , auquel cas  $\Gamma_z = \{\text{id}, ST, (ST)^2\}$ ,
- $z = -\bar{\omega} = e^{\frac{i\pi}{3}} = \omega + 1$ , auquel cas  $\Gamma_z = \{\text{id}, TS, (TS)^2\}$ .

- (2) Montrer que la projection canonique  $\mathbb{H}_{\mathbb{R}}^2 \rightarrow M = \Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$  est un revêtement ramifié d'ordre 2 au-dessus de  $\Gamma i$  et d'ordre 3 au-dessus de  $\Gamma \omega$  où  $\omega = e^{2i\pi/3}$ . Montrer que  $M$  est homéomorphe au disque ouvert unité.
- (3) Déterminer les orbites périodiques du flot géodésique sur  $T^1M = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ . On rappelle (voir par exemple [Khi]) que le développement en fraction continue d'un irrationnel quadratique est périodique à partir d'un certain rang.
- (4) Déterminer les orbites périodiques du flot horocyclique sur  $T^1M$ .
- (5) Montrer que la réunion des orbites périodiques du flot géodésique est dense dans  $T^1M$ . On rappelle (voir par exemple [Khi]) que si  $\alpha = [\overline{a_0, a_1, \dots, a_{m-1}, a_m}]$  est un irrationnel quadratique dans  $\mathbb{R}$  dont le développement en fraction continue est purement périodique de période  $a_0, a_1, \dots, a_{m-1}, a_m$  (donc de premier coefficient  $a_0 = \lfloor \alpha \rfloor \geq 1$ ), si  $\alpha^\sigma$  est le conjugué de Galois de  $\alpha$  (l'autre racine d'un polynôme quadratique à coefficients entiers dont  $\alpha$  est une racine), alors le développement en fraction continue de  $-\frac{1}{\alpha^\sigma}$  est  $[\overline{a_m, a_{m-1}, \dots, a_1, a_0}]$ , périodique de période l'image miroir  $a_m, a_{m-1}, \dots, a_1, a_0$  de la période de  $\alpha$ .
- (6) Montrer que pour tout  $v \in T^1\mathbb{H}_{\mathbb{R}}^2$ , l'orbite de  $\Gamma v \in T^1M$  par le flot géodésique de  $T^1M$  est d'ensemble  $\omega$ -limite non vide dans  $T^1M$  si et seulement si le point à l'infini du rayon géodésique de  $\mathbb{H}_{\mathbb{R}}^2$  défini par  $v$  n'appartient pas à  $\mathbb{Q} \cup \{\infty\}$ .
- (7) Notons  $\text{pr}_2 : T^1M \rightarrow \mathbb{S}_1$  l'application bien définie  $m_{\text{Liou}}$ -presque partout, telle que si  $\tilde{v} \in T^1\mathbb{H}_{\mathbb{R}}^2$ , alors  $\text{pr}_2(\Gamma \tilde{v}) = \xi$  avec  $(\Gamma \tilde{v}) \cap (\mathcal{F} \times \mathbb{S}_1) = \{(z, \xi)\}$ .<sup>197</sup> Montrer, en utilisant le théorème ergodique de Birkhoff 2.2, que pour  $m_{\text{Liou}}$ -presque tout  $v \in T^1M$ , le nombre d'enroulements moyen du rayon géodésique défini par  $v$  autour de la pointe de  $M$  est nul, c'est-à-dire que la limite

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \text{Re } \text{pr}_2(\mathfrak{g}^t v) dt$$

existe et vaut 0.

**Exercice E.56.** Notons  $\Gamma = \text{PSL}_2(\mathbb{Z})$  le groupe modulaire,  $M = \Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$  la courbe modulaire et  $T^1M = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ . Le but de cet exercice est de montrer, dans le cas particulier où  $\Gamma$  est le groupe modulaire, le *théorème de Dani*<sup>176</sup>-*Furstenberg* suivant : si  $\Gamma$  est un réseau de  $G = \text{PSL}_2(\mathbb{R})$ , les seules mesures boréliennes sur  $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ , de probabilité, invariantes par le flot horocyclique et qui sont ergodiques pour ce flot sont :

- la mesure de Liouville  $m_{\text{Liou}}$ , renormalisée pour être de probabilité,
- les mesures de Lebesgue  $\mathcal{L}_{\mathcal{O}}$  sur les orbites périodiques  $\mathcal{O}$  du flot horocyclique : si  $v \in T^1M$  est périodique de période  $T > 0$  par le flot horocyclique, alors la *mesure*

<sup>197.</sup> Autrement dit, pour tout élément  $v \in \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ , l'élément  $\text{pr}_2(v)$  est la direction  $\xi \in \mathbb{S}_1$  d'un relevé  $(z, \xi) \in T^1\mathbb{H}_{\mathbb{R}}^2$  de  $v$  par la projection canonique  $T^1\mathbb{H}_{\mathbb{R}}^2 \rightarrow \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  dont l'origine appartient à  $\mathcal{F}$ , ce relevé étant unique si  $z \in \mathring{\mathcal{F}}$ .

de Lebesgue  $\mathcal{L}_{\mathcal{O}(v)}$  sur  $\mathcal{O}(v)$  est la mesure image de la mesure de Lebesgue sur  $[0, T]$  par l'application  $s \mapsto \mathfrak{h}^s(v)$ , renormalisée pour être de probabilité.

De plus, si  $\Gamma$  est un réseau uniforme, alors la mesure de Liouville  $m_{\text{Liou}}$  est uniquement ergodique pour le flot horocyclique.

- (1) Montrer que ces mesures sont ergodiques pour le flot horocyclique.

Dans la suite, fixons une mesure borélienne  $\mu$  sur  $T^1M$ , de probabilité, invariante par le flot horocyclique et qui est ergodique.

- (2) Montrer que si  $v \in T^1\mathbb{H}_{\mathbb{R}}^2$  est tel que l'ensemble  $\omega$ -limite  $\omega_{\mathfrak{g}}(\Gamma v)$  de l'image de  $v$  dans  $T^1M$  pour le flot géodésique est vide, alors le point à l'infini  $v_+$  du rayon géodésique défini par  $v$  appartient à  $\mathbb{Q} \cup \{\infty\}$ .
- (3) Notons  $R$  l'ensemble des points  $v \in T^1M$  dont l'ensemble  $\omega$ -limite pour le flot géodésique est non vide. Montrer, en utilisant le théorème ergodique de Birkhoff 2.2, que si  $\mu(R) = 0$ , alors  $\mu$  est une mesure de Lebesgue sur une orbite périodique du flot horocyclique.
- (4) Montrer que si  $\mu(R) = 1$ , alors  $\mu$  est la mesure de Liouville, renormalisée pour être de probabilité. Pour cela, on pourra procéder de la manière suivante.

i) Munissons  $\mathbb{R}^2$  de la norme euclidienne usuelle,  $\mathcal{M}_2(\mathbb{R})$  de la norme d'opérateur associée,  $\text{SL}_2(\mathbb{R})$  de la distance définie par la formule (69) et  $G = \text{PSL}_2(\mathbb{R})$  de la distance quotient. Munissons  $T^1\mathbb{H}_{\mathbb{R}}^2$  de la distance telle que l'homéomorphisme (défini dans le lemme 9.15)  $\Phi : G \rightarrow T^1\mathbb{H}_{\mathbb{R}}^2$  soit une isométrie, et munissons  $T^1M$  de la distance quotient. Montrer qu'il existe  $\kappa > 0$  tel que pour tous les  $v \in T^1M$  et  $t, s \in \mathbb{R}$ , nous avons

$$d(v, \mathfrak{g}^t v) \leq \kappa |t|, \quad d(v, \mathfrak{h}^s v) \leq \kappa |s|, \quad d(v, \bar{\mathfrak{h}}^s v) \leq \kappa |s|,$$

et, si  $t \geq 0$ ,

$$d(\mathfrak{g}^{-t} v, \mathfrak{g}^{-t} \bar{\mathfrak{h}}^s v) \leq \kappa |s| e^{-t}, \quad d(\mathfrak{g}^t v, \mathfrak{g}^t \mathfrak{h}^s v) \leq \kappa |s| e^{-t}.$$

ii) Pour tous les  $f \in C_c^0(T^1M)$  et  $t \geq 1$ , posons

$$M_t f : v \mapsto \int_0^1 f(\mathfrak{g}^{-\ln t} \mathfrak{h}^s(v)) ds.$$

Montrer (ou admettre) que la famille  $(M_t f)_{t \in \mathbb{R}}$  d'applications de  $T^1M$  dans  $\mathbb{R}$  est équicontinue, et uniformément bornée.

iii) Pour tout  $f \in C_c^0(T^1M)$ , montrer qu'il existe une suite réelle strictement croissante  $(t_n)_{n \in \mathbb{N}}$  convergeant vers  $+\infty$  telle que

$$\lim_{n \rightarrow +\infty} M_{t_n} f \rightarrow \frac{1}{\|m_{\text{Liou}}\|} \int_{T^1M} f dm_{\text{Liou}}.$$

En utilisant le théorème ergodique de Birkhoff 2.2, en déduire que  $\mu = \frac{m_{\text{Liou}}}{\|m_{\text{Liou}}\|}$ .

- (5) Conclure.

## 9.7 Indications pour la résolution des exercices

**Correction de l'exercice E.54.** (1) Par conjugaison (en faisant attention à l'ordre : nous avons  $\mathfrak{g}^t \mathfrak{h}^s \mathfrak{g}^{-t} \Phi(g) = \Phi(g a_{-t} u_s a_t)$ ), la formule (83) découle du fait que

$$a_{-t} u_s a_t = \begin{bmatrix} e^{-t/2} & 0 \\ 0 & e^{t/2} \end{bmatrix} \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{bmatrix} = \begin{bmatrix} 1 & s e^{-t} \\ 0 & 1 \end{bmatrix} = u_s e^{-t}.$$

(2) Le fait que  $\bar{\mathfrak{h}}^s$  commute avec l'action de  $G$  pour tout  $s \in \mathbb{R}$  découle de la construction du flot horocyclique instable et du fait que par la formule (73), pour tout  $g \in G$  et  $(z, \xi) \in T^1 \mathbb{H}_{\mathbb{R}}^2$ , si  $g \cdot (z, \xi) = (z', \xi')$ , alors  $g \cdot (z, i\xi) = (z', i\xi')$ . En effet, si  $s \mapsto \bar{c}(s)$  est l'unique paramétrage par longueur d'arc hyperbolique de l'horocycle instable  $H^u(v)$  de  $v$  tel que  $v = (\bar{c}(0), i\bar{c}'(0))$ , alors  $s \mapsto g \cdot \bar{c}(s)$  est l'unique paramétrage par longueur d'arc hyperbolique de l'horocycle instable  $g \cdot H^u(v) = H^u(g \cdot v)$  de  $g \cdot v$  tel que (par la formule (73))  $g \cdot v = (g \cdot \bar{c}(0), g'(\bar{c}(0)) i \bar{c}'(0)) = (g \cdot \bar{c}(0), i(g \cdot \bar{c})'(0))$ . Donc

$$\bar{\mathfrak{h}}^s(g \cdot v) = (g \cdot \bar{c}(s), i(g \cdot \bar{c})'(s)) = (g \cdot \bar{c}(s), g'(\bar{c}(s)) i \bar{c}'(s)) = g \cdot \bar{\mathfrak{h}}^s(v).$$

Montrons que pour tous les  $g \in G$  et  $s \in \mathbb{R}$ , nous avons

$$\bar{\mathfrak{h}}^s(\Phi(g)) = \Phi(g u_s^-) \quad (84)$$

Puisque  $\Phi : G \rightarrow T^1 \mathbb{H}_{\mathbb{R}}^2$  est un  $C^\infty$ -difféomorphisme qui envoie la mesure de Haar de  $G$  (qui est invariante par translation à droite par  $u_s^-$ ) sur la mesure de Liouville de  $T^1 \mathbb{H}_{\mathbb{R}}^2$  (voir le lemme 9.15), ceci montrera que  $(\bar{\mathfrak{h}}^s)_{s \in \mathbb{R}}$  est un groupe à un paramètre de  $C^\infty$ -difféomorphismes de  $T^1 \mathbb{H}_{\mathbb{R}}^2$ , qui préserve la mesure de Liouville.

L'homographie  $z \mapsto -\frac{1}{z}$ , qui est une isométrie de  $\mathbb{H}_{\mathbb{R}}^2$ , envoie l'horocycle stable  $H^s(v)$  du vecteur tangent  $v = (i, i)$  sur l'horocycle instable  $H^u(v)$  du vecteur tangent  $v = (i, i)$  (voir la figure précédant le lemme 9.18). Comme  $s \mapsto i - s$  est un paramétrage par longueur d'arc hyperbolique de  $H^s(v)$ , l'application  $\bar{c} : s \mapsto -\frac{1}{i-s}$  est un paramétrage par longueur d'arc hyperbolique de  $H^u(v)$  tel que  $(\bar{c}(0), i\bar{c}'(0)) = (i, i \times 1) = v$ . Pour montrer la formule (84), il suffit, par équivariance, de vérifier que  $\bar{\mathfrak{h}}^s(i, i) = \Phi(u_s^-)$ . Par la définition du flot horocyclique instable, nous avons

$$\bar{\mathfrak{h}}^s(i, i) = (\bar{c}(s), i\bar{c}'(s)) = \left( -\frac{1}{i-s}, -i\frac{1}{(i-s)^2} \right).$$

Par la formule (73), nous avons

$$\Phi(u_s^-) = \begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix} \cdot (i, i) = \left( \frac{i}{is+1}, i\frac{1}{(is+1)^2} \right) = \left( -\frac{1}{i-s}, -i\frac{1}{(i-s)^2} \right).$$

Donc  $\bar{\mathfrak{h}}^s(i, i) = \Phi(u_s^-)$ , ce qu'il fallait établir.

Comme dans la question (1), par conjugaison (en faisant attention à l'ordre : nous avons  $\mathfrak{g}^t \bar{\mathfrak{h}}^s \mathfrak{g}^{-t} \Phi(g) = \Phi(g a_{-t} u_s^- a_t)$ ), la formule

$$\forall s, t \in \mathbb{R}, \quad \mathfrak{g}^t \circ \bar{\mathfrak{h}}^s \circ \mathfrak{g}^{-t} = \bar{\mathfrak{h}}^{s e^t} \quad (85)$$

découle du fait que

$$a_{-t} u_s^- a_t = \begin{bmatrix} e^{-t/2} & 0 \\ 0 & e^{t/2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix} \begin{bmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ s e^t & 1 \end{bmatrix} = u_s^- e^t.$$



Le fait que le flot horocyclique instable  $(\bar{\mathfrak{h}}^s)_{s \in \mathbb{R}}$  sur  $\Gamma \backslash T^1 \mathbb{H}_{\mathbb{R}}^2$  soit mélangeant si  $\Gamma$  est un réseau se démontre comme le corollaire 9.19.

(3) Soient  $\Gamma$  un réseau de  $\mathrm{PSL}_2(\mathbb{R})$ ,  $Y = \Gamma \backslash T^1 \mathbb{H}_{\mathbb{R}}^2$ , et  $f \in \mathbb{L}^2(Y, m_{\mathrm{Liou}})$  telle que  $f \circ \mathfrak{g}^t = f$  (presque partout). Soit  $s \in \mathbb{R}$ . Par l'invariance de la mesure de Liouville par le flot géodésique, par la formule (83) et par l'invariance de  $f$  par le flot géodésique, nous avons

$$\begin{aligned} \|f \circ \mathfrak{h}^s - f\|_{\mathbb{L}^2} &= \|f \circ \mathfrak{h}^s \circ \mathfrak{g}^{-t} - f \circ \mathfrak{g}^{-t}\|_{\mathbb{L}^2} = \|f \circ \mathfrak{g}^{-t} \circ \mathfrak{h}^{s e^{-t}} - f \circ \mathfrak{g}^{-t}\|_{\mathbb{L}^2} \\ &= \|f \circ \mathfrak{h}^{s e^{-t}} - f\|_{\mathbb{L}^2}, \end{aligned}$$

qui tend vers 0 quand  $t$  tend vers  $+\infty$ , par continuité.<sup>198</sup> Avec le même argument (en prenant  $t \rightarrow -\infty$ ) pour le flot horocyclique instable (en utilisant la formule (85)), nous avons donc que  $\|f \circ \mathfrak{h}^s - f\|_{\mathbb{L}^2} = 0$  et  $\|f \circ \bar{\mathfrak{h}}^{s'} - f\|_{\mathbb{L}^2} = 0$  pour tous les  $s, s' \in \mathbb{R}$ .

Par la commutativité des diagrammes précédant le corollaire 9.19 et par la formule (84), nous avons donc que l'application  $f \circ \Phi \in \mathbb{L}^2(\Gamma \backslash G, \mu_{\Gamma \backslash G})$  est invariante par la multiplication à droite par  $a_t, u_s, u_{s'}$  pour tous les  $t, s, s' \in \mathbb{R}$ . Puisque ces éléments engendrent  $G$ , la fonction  $f \circ \Phi$  est donc invariante par  $G$ , donc est constante presque partout par le corollaire 9.3. Par conséquent,  $f$  est constante presque partout, et le flot géodésique est ergodique.

**Correction de l'exercice E.55.** (1) Cette démonstration est extraite de [Ser, §VII.1.2].

Soient  $z \in \mathcal{F}$  et  $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma$  tels que  $\gamma \cdot z \in \mathcal{F}$ . Quitte à remplacer  $(z, \gamma)$  par  $(\gamma \cdot z, \gamma^{-1})$ , nous pouvons supposer que  $\mathrm{Im} z \geq \mathrm{Im} \gamma z$ , donc que  $|cz + d| \leq 1$  par la formule (72). Puisque  $\frac{\sqrt{3}}{2} |c| \leq |c \mathrm{Im} z| \leq |cz + d|$ , ceci implique que  $|c| \leq \frac{2}{\sqrt{3}}$ , donc  $c \in \{-1, 0, 1\}$ . Quitte à changer globalement les signes de  $a, b, c, d$ , nous pouvons supposer que  $c \geq 0$ .

Si  $c = 0$ , alors  $d = \pm 1$ , donc  $a = \frac{1}{d} = \pm 1$  et  $\gamma = T^{\pm b}$ . Comme les parties réelles de  $z$  et de  $\gamma \cdot z$  sont dans  $[-\frac{1}{2}, \frac{1}{2}]$ , ceci implique ou bien que  $b = 0$  (auquel cas  $\gamma = \mathrm{id}$  et  $z = \gamma \cdot z$ ) ou bien que  $b = \pm 1$  et  $\gamma \cdot z = z \pm 1$ , auquel cas l'un des réels  $\mathrm{Re} z$  et  $\mathrm{Re}(\gamma \cdot z)$  est égal à  $-1/2$  et l'autre à  $1/2$  (et en particulier ils sont distincts).

Supposons maintenant que  $c = 1$ . Si  $d = 0$ , alors  $|z| \leq 1$ , donc  $|z| = 1$  puisque  $z \in \mathcal{F}$ . Comme  $ad - bc = 1$ , nous en déduisons que  $b = -1$  et  $\gamma \cdot z = a - 1/z$ . Comme le cercle unité ne rencontre un de ses translatés par un entier non trivial qu'en  $\omega = e^{\frac{2i\pi}{3}}$  ou  $-\bar{\omega} = \omega + 1$ , nous obtenons  $a = 0$  (auquel cas  $\gamma = S$  et  $\gamma \cdot z = z$  si et seulement si  $z = i$ ) ou  $a = 1$  (auquel cas  $\gamma = (ST)^2$  et  $z = \gamma \cdot z = \omega$ ) ou  $a = -1$  (auquel cas  $\gamma = ST$ ,  $z = \omega$ , et  $\gamma \cdot z = -\bar{\omega}$ ), ce qui montre le résultat. Si  $d \neq 0$ , alors  $z$  appartient à l'intersection avec  $\mathcal{F}$  du disque fermé de centre  $-d$  et de rayon 1, donc ou bien  $d = 1$  et  $z = \omega$ , ou bien  $d = -1$  et  $z = -\bar{\omega}$ .

Si  $d = 1$  et  $z = \omega$ , alors  $a - b = 1$  et  $\gamma \cdot z = a - 1/(z + 1) = a + z$ , donc  $a = 0$  ou  $a = 1$ , ce qui implique le résultat. Le cas  $z = -\bar{\omega}$  et  $d = -1$  se traite de la même manière.

(2) Par la question (1), si  $\widetilde{M}' = \mathbb{H}_{\mathbb{R}}^2 - (\Gamma i \cup \Gamma \omega)$  est le demi-plan supérieur privé des orbites de  $i$  et de  $\omega$ , alors le groupe  $\Gamma$  agit proprement et librement sur  $\widetilde{M}'$ . Donc la projection canonique  $\widetilde{M}' \rightarrow \Gamma \backslash \widetilde{M}'$  est un revêtement, dont l'image est exactement le complémentaire dans  $M$  des points  $\Gamma i$  et  $\Gamma \omega$ . Puisque les stabilisateurs de  $i$  et  $\omega$  dans  $\Gamma$  sont des groupes cycliques d'ordre 2 et 3 respectivement, ceci démontre la première assertion de la question (2).

198. Cet argument est une variante du phénomène de Mautner décrit dans la partie 9.3.

Par la question (1), l'espace topologique quotient  $M$  est homéomorphe au recollement de deux copies de  $\{z \in \mathbb{C} : |z| \geq 1, 0 \leq \operatorname{Re} z \leq \frac{1}{2}\}$  par l'identité sur le bord. Donc  $M$  est homéomorphe au recollement de deux moitiés de disques ouverts par l'identité sur leur diamètre de bord, donc est homéomorphe à un disque.

(3) Nous appellerons *géodésique fermée* de  $M$  toute image par la projection canonique  $T^1M \rightarrow M$  (définie par  $\Gamma(z, \xi) \mapsto \Gamma \cdot z$ ) d'une orbite périodique du flot géodésique sur  $T^1M$ . Ce sont donc exactement les images par la projection canonique  $\mathbb{H}_{\mathbb{R}}^2 \rightarrow M = \Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$  des images  $\ell(\mathbb{R})$  des géodésiques  $\ell$  de  $\mathbb{H}_{\mathbb{R}}^2$  définies sur  $\mathbb{R}$  telles qu'il existe  $\kappa > 0$  et  $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  dans  $\Gamma$  tels que pour tout  $t \in \mathbb{R}$ , nous ayons  $\gamma \circ \ell(t) = \ell(t + \kappa)$ .

En particulier, un tel  $\gamma$  fixe les deux points (distincts) à l'infini de  $\ell$ , qui sont solutions de l'équation  $\frac{az+b}{cz+d} = z$ .

Les seuls éléments  $\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \in \Gamma$  qui fixent  $\infty$  sont ceux qui vérifient  $c' = 0$ , donc  $a = d = \pm 1$  puisque  $a, d \in \mathbb{Z}$  et  $ad - bc = 1$ . Ils agissent donc sur  $\mathbb{H}_{\mathbb{R}}^2$  par des translations entières  $z \mapsto z + k$  où  $k \in \mathbb{Z}$ , qui n'ont pas d'autre point fixe que  $\infty$  dans la sphère de Riemann  $\mathbb{C} \cup \{\infty\}$ .

Donc une telle géodésique  $\ell$  n'est pas une demi-droite verticale, donc est un arc de cercle perpendiculaire à l'axe des réels, et en particulier  $c \neq 0$  et  $cz + d \neq 0$  si  $z$  est une extrémité de  $\ell$ . Donc les points à l'infini de  $\ell$  sont les deux solutions de l'équation quadratique à coefficients entiers  $cz^2 + (d - a)z + b = 0$ . Ces points à l'infini ne sont pas rationnels, car s'ils étaient de la forme  $\frac{p}{q}$  avec  $p$  et  $q$  premier entre eux, alors par le théorème de Bezout, il existerait  $r, s \in \mathbb{Z}$  tels que  $pr - qs = 1$  et donc  $\gamma' = \begin{bmatrix} p & s \\ q & r \end{bmatrix}^{-1}$ , qui appartient à  $\Gamma$ , enverrait la géodésique périodique  $\ell$  sur une géodésique périodique qui est une demi-droite verticale, ce que nous avons exclu. Donc  $\ell$  admet pour couple d'extrémités un couple de nombres algébriques quadratiques  $\alpha, \alpha^\sigma$  conjugués de Galois (c'est-à-dire solutions distinctes d'une équation quadratique à coefficients entiers).

Réciproquement, soit  $\alpha$  un irrationnel quadratique, et montrons qu'il existe un élément  $\gamma \in \Gamma$  qui agit par une translation non triviale sur la géodésique  $\ell$  définie sur  $\mathbb{R}$  dont les points à l'infinis sont  $\alpha$  et son conjugué de Galois  $\alpha^\sigma$ , donc se projette sur une géodésique fermée de  $M$ . Quitte à traduire  $\alpha$  par un entier (et à conjuguer l'élément cherché  $\gamma$  par un élément  $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$  où  $k \in \mathbb{Z}$ ), nous pouvons supposer que  $\alpha \in ]0, 1[$ . Par un résultat

classique, son développement en fraction continue  $\alpha = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots}}}$ , où  $a_i \in \mathbb{N} - \{0\}$ ,

est périodique à partir d'un certain rang. Notons que  $x = \frac{1}{a_1 + \frac{1}{a_2 + y}}$  si et seulement si

$x = \gamma \cdot y$ , où  $\gamma = \begin{bmatrix} 1 & a_2 \\ a_1 & 1 + a_1 a_2 \end{bmatrix} \in \Gamma$ . Quitte à modifier  $\alpha$  par une homographie entière (et à conjuguer l'élément cherché  $\gamma$  par la matrice de cette homographie), nous pouvons donc supposer que le développement en fraction continue de  $\alpha$  est périodique à partir du rang

1. Par conséquent,  $\alpha$  est solution d'une équation  $\alpha = \frac{1}{a_1 + \frac{1}{a_2 + \cdots + \frac{1}{a_{2n} + \alpha}}}$ , donc est un

point fixe de l'homographie (à coefficients strictement positifs) associée à un élément non trivial  $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma$ .

Cet élément  $\gamma$  fixe aussi le conjugué de Galois  $\alpha^\sigma$ , car  $(\frac{a\alpha+b}{c\alpha+d})^\sigma = \frac{a\alpha^\sigma+b}{c\alpha^\sigma+d}$ . Puisque  $\gamma$  agit par isométrie pour la distance hyperbolique de  $\mathbb{H}_{\mathbb{R}}^2$ , et puisque les seules isométries de  $\mathbb{R}$  fixant  $-\infty$  et  $+\infty$  sont les translations, il existe  $\kappa \in \mathbb{R}$  tel que pour tout  $t \in \mathbb{R}$ , nous ayons  $\gamma \circ \ell(t) = \ell(t + \kappa)$ .

Le seul élément de  $G$  fixant point par point l'axe imaginaire dans  $\mathbb{H}_{\mathbb{R}}^2$ , donc fixant  $i$  et les deux points à l'infini  $0, \infty$ , est l'identité. En effet, pour tous les  $a', b', c', d' \in \mathbb{R}$  tels que  $a'd' - b'c' = 1$ , si  $\frac{b'}{d'} = 0$ ,  $\frac{a'}{c'} = \infty$  et  $\frac{a'i+b'}{c'i+d'} = i$ , alors  $b' = c' = 0$  et  $\frac{a'}{d'} = a'd' = 1$ , donc  $a' = d' = \pm 1$ . Par conjugaison, la seule isométrie de  $\mathbb{H}_{\mathbb{R}}^2$  fixant point par point une géodésique est donc l'identité. Donc quitte à remplacer  $\gamma$  par  $\gamma^{-1}$ , nous pouvons supposer que  $\kappa > 0$ . Ceci montre l'existence d'un élément  $\gamma$  cherché.

Les orbites périodiques du flot géodésiques sont donc les images par la projection canonique  $T^1\mathbb{H}_{\mathbb{R}}^2 \rightarrow T^1M$  des vecteurs unitaires tangents à une géodésique de  $\mathbb{H}_{\mathbb{R}}^2$  définie sur  $\mathbb{R}$  dont les points à l'infini sont deux irrationnels quadratiques conjugués de Galois.

(4) Les orbites périodiques du flot horocyclique sur  $T^1M$  sont les images par la projection canonique  $T^1\mathbb{H}_{\mathbb{R}}^2 \rightarrow T^1M$  des orbites du flot horocyclique sur  $T^1\mathbb{H}_{\mathbb{R}}^2$  qui sont invariantes (modulo reparamétrage par translation) par un élément non trivial  $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  dans  $\Gamma$ . Cet élément va donc avoir comme seul point fixe à l'infini le point à l'infini de l'horocycle projection de l'orbite du flot horocyclique par la projection  $T^1\mathbb{H}_{\mathbb{R}}^2 \rightarrow \mathbb{H}_{\mathbb{R}}^2$ . Puisque  $a, b, c, d$  sont entiers, si l'équation  $\frac{az+b}{cz+d} = z$  admet une unique solution, alors cette solution appartient à  $\mathbb{Q} \cup \{\infty\}$ .

Réciproquement, montrons que tout horocycle  $\mathcal{H}$  centré en un point de  $\mathbb{Q} \cup \{\infty\}$  est la projection par  $T^1\mathbb{H}_{\mathbb{R}}^2 \rightarrow \mathbb{H}_{\mathbb{R}}^2$  d'une orbite périodique modulo  $\Gamma$  du flot horocyclique.

Le groupe  $\Gamma$  agit transitivement sur  $\mathbb{Q} \cup \{\infty\}$ , par l'argument de Bezout vu ci-dessus : pour tout  $\frac{p}{q} \in \mathbb{Q}$  avec  $p$  et  $q$  premiers entre eux et  $q \neq 0$ , si  $r, s \in \mathbb{Z}$  sont tels que  $pr - qs = 1$ , alors l'image de  $\infty$  par  $\begin{bmatrix} p & s \\ q & r \end{bmatrix}$  est exactement  $\frac{p}{q}$ . Puisque le flot horocyclique commute avec l'action de  $\Gamma$ , nous pouvons donc supposer que le point à l'infini de  $\mathcal{H}$  est  $\infty$ . Mais alors l'élément  $u_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$  est non trivial, et il préserve toute horosphère centrée en  $\infty$  (puisque son homographie associée est  $z \mapsto z + 1$ ), donc l'orbite du flot horocyclique correspondante dans  $\Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$  est périodique.

(5) Soient  $x, y \in \mathbb{R}$  deux réels distincts. Montrons qu'il existe une suite  $(\alpha_n)_{n \in \mathbb{N}}$  d'irrationnels quadratiques tels que  $\lim_{n \rightarrow +\infty} \alpha_n = x$  and  $\lim_{n \rightarrow +\infty} \alpha_n^\sigma = y$ , ce qui implique le résultat par la question (3). Quitte à modifier  $x$  et  $y$  par une même homographie rationnelle et à remplacer  $x$  et  $y$  par des points arbitrairement proches, nous pouvons supposer que  $x > 1$  et  $y \in ]-1, 0[$ . Notons  $(a_n)_{n \in \mathbb{N}}$  les coefficients du développement en fraction continue de  $x$  et  $(b_n)_{n \in \mathbb{N}}$  les coefficients du développement en fraction continue de  $-\frac{1}{y}$ , qui vérifient donc

$a_0 \geq 1$  et  $b_0 \geq 1$ . Pour tout  $n \in \mathbb{N}$ , posons  $\alpha_n = \overline{[a_0, a_1, \dots, a_n, b_n, \dots, b_1, b_0]}$ . Alors la suite  $(\alpha_n)_{n \in \mathbb{N}}$  convient, par la propriété appelée des fractions continues.

(6) Si  $\mathcal{H}$  est un horocycle, définissons l'horoboule (ouverte)  $\text{HB}$  de bord  $\mathcal{H}$  comme le demi-plan ouvert au-dessus de  $\mathcal{H}$ , si  $\mathcal{H}$  est une droite horizontale, ou comme la boule ouverte euclidienne dont le bord dans le plan euclidien est la réunion de  $\mathcal{H}$  et son point à l'infini, si  $\mathcal{H}$  est un cercle tangent à l'axe réel (privé du point de tangence).

Par la description des géodésiques et des horocycles, tout rayon géodésique qui rentre dans l'intérieur d'une horoboule, et qui ne converge pas vers son point à l'infini, en ressort. Par la description du domaine fondamental de  $\Gamma$ , si  $\text{HB}_\infty = \{z \in \mathbb{C} : \text{Im } z \geq 1\}$ , alors l'image de  $\text{HB}_\infty$  par un élément  $\gamma \in \Gamma$  est ou bien égale à  $\text{HB}_\infty$ , ou bien d'intérieur disjoint de  $\text{HB}_\infty$ . Nous avons vu qu'un élément de  $\mathbb{R} \cup \{\infty\}$  est le point à l'infini d'une horoboule  $\gamma \text{HB}_\infty$  pour  $\gamma \in \Gamma$  si et seulement s'il appartient à  $\mathbb{Q} \cup \{\infty\}$ . Toujours par la description du domaine fondamental de  $\Gamma$ , le complémentaire  $C$  dans  $\mathbb{H}_{\mathbb{R}}^2$  de l'intérieur de  $\bigcup_{\gamma \in \Gamma} \gamma \text{HB}_\infty$  est de quotient compact par  $\Gamma$  (car  $\mathcal{F}$  privé des points intérieurs à  $\text{HB}_\infty$  est compact). Donc tout rayon géodésique  $\rho$  dans  $\mathbb{H}_{\mathbb{R}}^2$  ou bien converge vers le point à l'infini de  $\gamma \text{HB}_\infty$  pour un certain  $\gamma \in \Gamma$ , ou bien revient une infinité de fois dans  $C$ . Par conséquent, l'image dans  $T^1M$  du vecteur tangent à l'origine de  $\rho$  est d'ensemble  $\omega$ -limite non vide pour le flot géodésique dans  $T^1M$  si et seulement si le point à l'infini de  $\rho$  n'appartient pas à  $\mathbb{Q} \cup \{\infty\}$ .

(7) Puisque la mesure de Liouville  $m_{\text{Liou}}$  sur  $T^1M$  est finie (par la proposition 9.16), puisque le flot géodésique sur  $T^1M$  est ergodique (voir le corollaire 9.20, ainsi que l'exercice E.54 (3)), puisque l'application  $\text{Re} \circ \text{pr}_2$  est mesurable bornée, donc intégrable pour  $m_{\text{Liou}}$ , le théorème ergodique de Birkhoff 2.2 dit que pour  $m_{\text{Liou}}$ -presque tout  $v \in T^1M$ , la limite suivante existe et a pour valeur celle indiquée, qui se simplifie en utilisant la formule (75),

$$\begin{aligned} \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \text{Re} \circ \text{pr}_2(\mathfrak{g}^t v) dt &= \frac{1}{\|m_{\text{Liou}}\|} \int_{T^1M} \text{Re} \circ \text{pr}_2 dm_{\text{Liou}} \\ &= \int_{z \in \mathcal{F}} \left( \int_{\theta \in \mathbb{R}/\mathbb{Z}} \text{Re}(e^{2i\pi\theta}) d\theta \right) \frac{dx dy}{y^2} = 0. \end{aligned}$$

**Correction de l'exercice E.56.** (1) Le fait que le flot horocyclique pour la mesure de Liouville  $m_{\text{Liou}}$  est ergodique a été vu dans le corollaire 9.20, puisque  $\Gamma$  est un réseau (par la proposition 9.16).

Le flot horocyclique agissant sur une orbite périodique de période  $T > 0$  est conjugué au groupe à un paramètre des translations du cercle  $\mathbb{R}/T\mathbb{Z}$ . L'ergodicité (et même l'unique ergodicité) des mesures de Lebesgue sur les orbites périodiques du flot horocyclique découle donc du fait que la mesure de Lebesgue sur le cercle  $\mathbb{R}/T\mathbb{Z}$  est l'unique mesure (borélienne positive finie) invariante par les translations, modulo multiplication par un scalaire strictement positif.

(2) Ceci découle de la question (6) de l'exercice E.55.

(3) L'ensemble  $R$  est mesurable. Puisque nous supposons que  $\mu(R) = 0$ , nous avons  $\mu({}^c R) = 1$ . Notons  $A$  une partie dénombrable dense de l'espace des fonctions continues à support compact sur  $T^1M$  pour la norme uniforme. Par le théorème ergodique de Birkhoff 2.2 et par dénombrabilité, il existe  $v \in {}^c R$  tel que, pour tout  $f \in A$ , nous avons

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t f(\mathfrak{h}^s(v)) ds = \mu(f).$$

Puisque  $v$  est l'image d'un vecteur  $\tilde{v} \in T^1\mathbb{H}_{\mathbb{R}}^2$  tel que le point à l'infini du rayon géodésique défini par  $\tilde{v}$  appartient à  $\mathbb{Q} \cup \{\infty\}$  par la question (2), et par la question (4) de l'exercice E.55, l'orbite de  $v$  par le flot horocyclique est périodique, de période  $T > 0$ . Nous avons donc, pour tout  $f \in A$ ,

$$\mu(f) = \lim_{n \rightarrow +\infty} \frac{1}{nT} \int_0^{nT} f(\mathfrak{h}^s(v)) ds = \frac{1}{T} \int_0^T f(\mathfrak{h}^s(v)) ds .$$

Par densité de  $A$  dans  $C_c^0(T^1M)$ , ceci est vrai pour tout  $f \in C_c^0(T^1M)$ , et donc  $\mu$  est la mesure de Lebesgue sur une orbite périodique du flot horocyclique.

(4) i) Soient  $t, s \in \mathbb{R}$ . Munissons  $\mathcal{M}_2(\mathbb{R})$  de sa norme d'opérateur sur l'espace euclidien usuel  $\mathbb{R}^2$ . Remarquons que

$$\|a_t - \text{id}\| \leq e^{\frac{|t|}{2}} - 1, \quad \|u_s - \text{id}\| \leq |s| \quad \text{et} \quad \|u_s^- - \text{id}\| \leq |s| .$$

Pour tout  $g \in G = \text{PSL}_2(\mathbb{R})$ , nous avons par la formule (82), puisque  $\Phi$  est une isométrie et puisque la distance sur  $G$  est invariante à gauche,

$$\begin{aligned} d(\Phi(g), \mathfrak{g}^t \Phi(g)) &= d(\Phi(g), \Phi(g a_t)) = d(g, g a_t) = d(e, a_t) \\ &\leq \ln(1 + (e^{\frac{|t|}{2}} - 1)) + \ln(1 + (e^{\frac{|t|}{2}} - 1)) = |t| . \end{aligned}$$

De même (et le calcul pour le flot horocyclique instable est identique), nous avons

$$d(\Phi(g), \mathfrak{h}^s \Phi(g)) = d(\Phi(g), \Phi(g u_s)) = d(g, g u_s) = d(e, u_s) \leq 2 \ln(1 + |s|) \leq 2|s| .$$

Si  $t \in \mathbb{R}$ , alors par la formule (83) (et le calcul pour le flot horocyclique instable est analogue en utilisant la formule (85)), nous avons

$$\begin{aligned} d(\mathfrak{g}^t \Phi(g), \mathfrak{g}^t \mathfrak{h}^s \Phi(g)) &= d(\mathfrak{g}^t \Phi(g), \mathfrak{h}^{s e^{-t}} \mathfrak{g}^t \Phi(g)) = d(\Phi(g a_t), \Phi(g a_t u_{s e^{-t}})) \\ &= d(g a_t, g a_t u_{s e^{-t}}) = d(e, u_{s e^{-t}}) \leq 2|s| e^{-t} . \end{aligned}$$

Le premier point en découle avec  $\kappa = 2$ .

ii) Soit  $f \in C_c^0(T^1M)$ . Nous avons  $\|M_t f\|_{\infty} \leq \|f\|_{\infty}$  pour tout  $t \geq 1$ , donc la famille  $(M_t f)_{t \in \mathbb{R}}$  est uniformément bornée.

Soient  $\epsilon > 0$  et  $v_0 \in T^1M$ . Par l'uniforme continuité de  $f$ , il existe  $\delta > 0$  tel que si  $d(v, v') \leq 2\kappa\delta$ , alors  $|f(v) - f(v')| \leq \epsilon$ . Pour tous les  $v \in T^1M$  et  $t \in [1, +\infty[$ , posons

$$K_v = \{\mathfrak{g}^{t'} \bar{\mathfrak{h}}^r \mathfrak{h}^s(v) : (t', r, s) \in [-\delta, \delta]^2 \times [0, 1]\} ,$$

et

$$N_t f(v) = \frac{1}{m_{\text{Liou}}(K_v)} \int_{K_v} f \circ \mathfrak{g}^{-\ln t} dm_{\text{Liou}}$$

la moyenne de  $f \circ \mathfrak{g}^{-\ln t}$  sur  $K_v$ .

Puisque  $u_s u_r^- a_{t'} = \begin{bmatrix} e^{\frac{t'}{2}}(1+rs) & s e^{-\frac{t'}{2}} \\ r e^{\frac{t'}{2}} & e^{-\frac{t'}{2}} \end{bmatrix}$  et puisque  $G$  agit simplement transitivement sur  $T^1\mathbb{H}_{\mathbb{R}}^2$ , pour tout  $\tilde{v} \in T^1\mathbb{H}_{\mathbb{R}}^2$ , l'application  $\varphi$  du compact  $[-\delta, \delta]^2 \times [0, 1]$  dans  $T^1\mathbb{H}_{\mathbb{R}}^2$  définie par  $(t', r, s) \mapsto \mathfrak{g}^{t'} \circ \bar{\mathfrak{h}}^r \circ \mathfrak{h}^s(\tilde{v})$ , qui est continue et injective, est un homéomorphisme

sur son image. En particulier,  $K_v$  est un compact de  $T^1M$  d'intérieur non vide, et  $N_t f$  est bien définie. De plus,  $\varphi$  est  $C^\infty$  et envoie la mesure de Lebesgue sur  $[-\delta, \delta]^2 \times [0, 1]$  sur une mesure absolument continue par rapport à la mesure de Liouville, ayant une dérivée de Radon-Nikodym de classe  $C^\infty$ . Donc sur  $K_v$ , par l'invariance de  $dm_{\text{Liou}}$  par le flot horocyclique et par l'invariance de la mesure de Lebesgue de  $\mathbb{R}$  par translations, nous avons  $dm_{\text{Liou}} = \rho(t', r) dt' dr ds$  avec  $\rho : [-\delta, \delta]^2 \rightarrow ]0, +\infty[$  continue, et en particulier  $m_{\text{Liou}}(K_v) = \int_{[-\delta, \delta]^2} \rho(t', r) dt' dr$ . D'où, pour tous les  $v \in T^1M$  et  $t \in [1, +\infty[$ , nous avons

$$\begin{aligned} & m_{\text{Liou}}(K_v)(M_t f(v) - N_t f(v)) \\ &= \int_0^1 \left( \int_{[-\delta, \delta]^2} (f \circ \mathfrak{g}^{-\ln t} \circ \mathfrak{h}^s(v) - f \circ \mathfrak{g}^{-\ln t}(\mathfrak{g}^{t'} \circ \bar{\mathfrak{h}}^r \circ \mathfrak{h}^s(v))) \rho(t', r) dt' dr \right) ds . \end{aligned}$$

En omettant les parenthèses dans ce qui suit, posons  $v' = \mathfrak{h}^s v$  et  $v'' = \mathfrak{g}^{-\ln t} \bar{\mathfrak{h}}^r v'$ . Par l'inégalité triangulaire et par l'assertion i), nous avons

$$\begin{aligned} & d(\mathfrak{g}^{-\ln t} \mathfrak{h}^s v, \mathfrak{g}^{-\ln t} \mathfrak{g}^{t'} \bar{\mathfrak{h}}^r \mathfrak{h}^s v) \\ & \leq d(\mathfrak{g}^{-\ln t} \mathfrak{h}^s v, \mathfrak{g}^{-\ln t} \bar{\mathfrak{h}}^r \mathfrak{h}^s v) + d(\mathfrak{g}^{-\ln t} \bar{\mathfrak{h}}^r \mathfrak{h}^s v, \mathfrak{g}^{-\ln t} \mathfrak{g}^{t'} \bar{\mathfrak{h}}^r \mathfrak{h}^s v) \\ & \leq d(\mathfrak{g}^{-\ln t} v', \mathfrak{g}^{-\ln t} \bar{\mathfrak{h}}^r v') + d(v'', \mathfrak{g}^{t'} v'') \leq \kappa |r| e^{-\ln t} + \kappa |t'| \leq 2\kappa \delta . \end{aligned}$$

Donc par l'uniforme continuité de  $f$ ,

$$m_{\text{Liou}}(K_v) |M_t f(v) - N_t f(v)| \leq \int_0^1 \int_{[-\delta, \delta]^2} \epsilon \rho(t', r) dt' dr = \epsilon m_{\text{Liou}}(K_v) .$$

D'où, par l'inégalité de Cauchy-Schwarz et par l'invariance de la mesure de Liouville par le flot géodésique,

$$\begin{aligned} |M_t f(v) - M_t f(v_0)| & \leq 2\epsilon + |N_t f(v) - N_t f(v_0)| \\ & \leq 2\epsilon + \int_{T^1M} \left| f \circ \mathfrak{g}^{-\ln t} \left( \frac{\mathbb{1}_{K_v}}{m_{\text{Liou}}(K_v)} - \frac{\mathbb{1}_{K_{v_0}}}{m_{\text{Liou}}(K_{v_0})} \right) \right| dm_{\text{Liou}} \\ & \leq 2\epsilon + \|f\|_2 \left\| \frac{\mathbb{1}_{K_v}}{m_{\text{Liou}}(K_v)} - \frac{\mathbb{1}_{K_{v_0}}}{m_{\text{Liou}}(K_{v_0})} \right\|_2 . \end{aligned}$$

Puisque la mesure de la frontière de  $K_v$  est nulle, la seconde norme  $\mathbb{L}^2$  (qui est indépendante de  $t$ ) tends vers 0 quand  $v$  tend vers  $v_0$ . Ceci montre l'équicontinuité de la famille  $(M_t f)_{t \in \mathbb{R}}$ .

iii) Soient  $f \in C_c^0(T^1M)$ . Pour tous les  $t > 0$  et  $v' \in T^1M$ , posons

$$H_t f(v') = \frac{1}{t} \int_0^t f(\mathfrak{h}^s(v')) ds ,$$

qui est l'intégrale de Birkhoff de  $f$  au temps  $t$  le long de l'orbite de  $v'$  par le flot horocyclique (stable). Par l'ergodicité de  $\mu$  et le théorème ergodique de Birkhoff, il existe une partie mesurable  $E$  de mesure pleine pour  $\mu$  telle que pour tout  $v \in E$ , nous ayons

$$\lim_{n \rightarrow +\infty} H_n f(v) = \int_{T^1M} f d\mu . \quad (86)$$

Fixons  $v \in R \cap E$ , ce qui est possible car  $R$  et  $E$  sont tous deux de mesure pleine pour  $\mu$  par l'hypothèse de la question (4). Par la question (6) de l'exercice E.55, l'ensemble  $\omega$ -limite

du vecteur tangent  $v$  est non vide. Il existe donc  $w \in T^1M$  et une suite réelle strictement croissante  $(t_n)_{n \in \mathbb{N}}$  convergeant vers  $+\infty$  telle que la suite  $(\mathbf{g}^{\ln t_n}(v))_{n \in \mathbb{N}}$  soit convergente vers  $w$ . En particulier l'adhérence  $K_v$  de  $\{\mathbf{g}^{\ln t_n}(v) : n \in \mathbb{N}\}$  est un compact de  $T^1M$ . Par la question ii) et le théorème d'Arzela-Ascoli, quitte à extraire, il existe une fonction continue  $F : T^1M \rightarrow \mathbb{C}$  telle que nous ayons  $F = \lim_{n \rightarrow +\infty} M_{t_n} f$  pour la convergence uniforme sur les compacts, donc pour la convergence dans  $\mathbb{L}^2(m_{\text{Liou}})$ .

Par la formule (83), pour tout  $v' \in T^1M$ , nous avons

$$\begin{aligned} H_{t_n} f(v') &= \frac{1}{t_n} \int_0^{t_n} f(\mathfrak{h}^s v') ds = \int_0^1 f(\mathfrak{h}^{st_n} v') ds = \int_0^1 f(\mathbf{g}^{-\ln t_n} \mathfrak{h}^s \mathbf{g}^{\ln t_n} v') ds \\ &= M_{t_n} f(\mathbf{g}^{\ln t_n} v'). \end{aligned} \quad (87)$$

Par l'ergodicité de la mesure de Liouville (qui a été renormalisée pour être de probabilité) par le flot horocyclique (voir le corollaire 9.20), pour  $m_{\text{Liou}}$ -presque tout  $v' \in T^1M$ , nous avons

$$\lim_{n \rightarrow +\infty} H_{t_n} f(v') = \int_{T^1M} f dm_{\text{Liou}}. \quad (88)$$

Puisque  $\lim_{n \rightarrow +\infty} \|M_{t_n} f - F\|_{\mathbb{L}^2(m_{\text{Liou}})} = 0$ , par la formule (87) et par l'invariance de la mesure de Liouville par le flot géodésique, nous avons donc  $\lim_{n \rightarrow +\infty} \|H_{t_n} f - F \circ \mathbf{g}^{-\ln t_n}\|_{\mathbb{L}^2(m_{\text{Liou}})} = 0$ . Par la formule (88), de nouveau par l'invariance de la mesure de Liouville par le flot géodésique, nous avons donc que  $F$  est  $m_{\text{Liou}}$ -presque partout constante égale à  $\int_{T^1M} f dm_{\text{Liou}}$ . Mais comme  $F$  est continue et puisque la mesure de Liouville est de support total, la fonction  $F$  est donc constante (partout), égale à  $\int_{T^1M} f dm_{\text{Liou}}$ . Par les formules (86) et (87), et par la convergence uniforme de  $M_{t_n} f$  vers  $F$  sur le compact  $K_v$ , nous avons donc

$$\int_{T^1M} f d\mu = \lim_{n \rightarrow +\infty} H_{t_n} f(v) = \lim_{n \rightarrow +\infty} M_{t_n} f(\mathbf{g}^{\ln t_n} v) = F(w) = \int_{T^1M} f dm_{\text{Liou}}.$$

Ceci étant vrai pour tout  $f \in C_c^0(T^1M)$ , nous avons  $\mu = m_{\text{Liou}}$ .

(5) Pour tout  $\eta \in \mathbb{R} \cup \{\infty\}$ , le flot horocyclique sur  $T^1M$  préserve l'ensemble des vecteurs unitaires tangents  $v$  tels que le point à l'infini du rayon géodésique défini par  $v$  est égal à  $\eta$ . Donc par la question précédente, l'ensemble mesurable  $R$  est invariant par le flot horocyclique. Puisque  $\mu$  est ergodique, nous avons donc  $\mu(R) = 0$  ou  $\mu(R) = 1$ . Donc les deux questions précédentes concluent.

## 10 Systèmes dynamiques hyperboliques

Les références principales pour ce chapitre sont [Sma, Bow3, Shu, Yoc2] et nous suivrons surtout cette dernière référence.

Dans ce chapitre, nous introduisons des systèmes dynamiques (à temps discret) très chaotiques, qui vérifient des hypothèses d'hyperbolicité (de contraction et dilatation dans des directions transverses). Nous montrons des phénomènes de stabilité sous petite perturbation pour ces systèmes dynamiques. Dit en terme de physiciens, il s'agit de montrer que "le chaos est stable par petite perturbation". Nous montrerons trois tels résultats de stabilité par petite perturbation, un cas linéaire (la proposition 10.3), un cas local (le

théorème de Grobman-Hartman 10.6) et un cas global (le théorème 10.8 de stabilité structurelle d’Anosov<sup>199</sup> pour les automorphismes linéaires du tore dits d’Anosov). Le plus gros travail de ce chapitre sera le théorème d’existence de variétés stables et instables (voir la partie 10.8).

Outre l’exemple du fer à cheval de Smale<sup>199</sup> traité dans la partie 10.5, les exemples principaux sont encore les automorphismes linéaires du tore, qui, lorsque leur matrice entière associée n’a pas de valeur propre de module 1, sont des archétypes des transformations “chaotiques”. Les démonstrations de ce chapitre reposent surtout sur le théorème du point fixe pour les applications contractantes.

## 10.1 Endomorphismes linéaires hyperboliques

Soient  $E$  un espace de Banach réel et  $T$  un endomorphisme linéaire continu<sup>200</sup> de  $E$ . Pour tout espace de Banach réel ou complexe  $E'$ , nous munirons l’espace  $\mathcal{L}(E')$  des endomorphismes linéaires continus de  $E'$  de la norme d’opérateur.<sup>201</sup>

L’endomorphisme linéaire continu  $T$  est dit *hyperbolique* s’il existe des sous-espaces vectoriels  $E^s$  et  $E^u$  de  $E$ , appelés les sous-espaces *stable* et *instable* de  $T$  respectivement, et  $\kappa'_s, \kappa'_u \in \mathbb{R}$  tels que  $0 < \kappa'_s < 1 < \kappa'_u$ , de sorte que

- (1) l’espace vectoriel  $E$  est somme directe de  $E^s$  et  $E^u$ , et la décomposition  $E = E^s \oplus E^u$  est dite *associée* à  $T$ ,
- (2) les sous-espaces stable et instable sont *invariants* par  $T$ , c’est-à-dire qu’ils vérifient  $T(E^s) \subset E^s$  et  $T(E^u) \subset E^u$ , et l’endomorphisme linéaire continu  $T|_{E^u}: E^u \rightarrow E^u$  est inversible,
- (3) il existe  $c > 0$  tel que

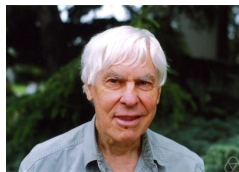
$$\forall n \in \mathbb{N}, \quad \|T|_{E^s}^n\| \leq c \kappa'_s{}^n, \quad \|T|_{E^u}^{-n}\| \leq c \kappa'_u{}^{-n}.$$

Si  $\kappa_s, \kappa_u \in \mathbb{R}$ , nous dirons que  $T$  est  $(\kappa_s, \kappa_u)$ -*hyperbolique* s’il existe des sous-espaces vectoriels  $E^s$  et  $E^u$  de  $E$  et  $\kappa'_s, \kappa'_u \in \mathbb{R}$  tels que  $0 < \kappa'_s < \kappa_s < 1 < \kappa_u < \kappa'_u$  et les conditions (1) à (3) ci-dessus soient vérifiées.

Nous verrons dans le lemme 10.1 ci-dessous que les sous-espaces vectoriels  $E^s$  et  $E^u$  sont alors fermés, et qu’une décomposition  $E = E^s \oplus E^u$  associée à  $T$  est alors unique. Lorsque  $E$  est de dimension finie, la définition ne dépend pas de la norme de  $E$ . Le passage de  $(\kappa'_s, \kappa'_u)$  à  $(\kappa_s, \kappa_u)$  dans la définition d’endomorphisme  $(\kappa_s, \kappa_u)$ -hyperbolique est mineure



Dimitry Anosov  
199. 1936-2014



Stephen Smale  
1930-

200. En dimension finie, la condition de continuité est automatique, mais il est crucial de ne pas l’oublier en dimension infinie. Nous aurons besoin du cadre de la dimension infinie dans des démonstrations ultérieures.

201. Avec les conventions usuelles si  $E'$  est nul, pour tout élément  $f \in \mathcal{L}(E')$ , rappelons que la *norme d’opérateur* de  $f$  est définie par  $\|f\| = \sup_{x \in E' - \{0\}} \frac{\|f(x)\|}{\|x\|}$ . Nous munirons  $\mathcal{L}(E')$  de cette norme. Notons que  $\mathcal{L}(E')$  est alors un espace de Banach réel ou complexe, et que le sous-espace  $\mathcal{GL}(E')$  des automorphismes linéaires continus de  $E'$  est un ouvert de  $\mathcal{L}(E')$  (voir par exemple [Pau6, §1.1]).



et purement technique (en particulier utilisé pour la proposition 10.2 en dimension infinie). L'important est l'idée de contraction / dilatation uniforme des espaces stables / instables.

Une norme  $\| \cdot \|'$  sur  $E$  est dite *adaptée* à un endomorphisme  $(\kappa_s, \kappa_u)$ -hyperbolique  $T$  si elle est équivalente à la norme de  $E$  (ce qui est automatique en dimension finie) et si

$$\forall x_s \in E^s, \forall x_u \in E^u, \quad \|x_s + x_u\|' = \max\{\|x_s\|', \|x_u\|'\}$$

et

$$\|T|_{E^s}\|' \leq \kappa_s, \quad \|T|_{E^u}^{-1}\|' \leq \kappa_u^{-1}.$$

Nous appellerons alors *constante d'hyperbolicité* de  $T$  la constante, qui dépend du choix de la norme adaptée  $\| \cdot \|'$ , définie par

$$\text{ch}(T) = \max\{\|T|_{E^s}\|', \|T|_{E^u}^{-1}\|'\} \leq \max\{\kappa_s, \kappa_u^{-1}\} < 1. \quad (89)$$

**Lemme 10.1.** *Reprenons les notations de la définition d'un endomorphisme linéaire hyperbolique  $T$  de  $E$ .*

- (1) *Une décomposition  $E = E^s \oplus E^u$  associée à  $T$  est unique.*
- (2) *Les sous-espaces vectoriels stable  $E^s$  et instable  $E^u$  sont fermés.*
- (3) *Il existe au moins une norme sur  $E$  adaptée à  $T$ .*

**Démonstration.** Soient  $0 < \kappa_s < 1 < \kappa_u < +\infty$  tels que  $T$  est  $(\kappa_s, \kappa_u)$ -hyperbolique. Soit  $\lambda = \max\{\kappa_s, \kappa_u^{-1}\} < 1$ , et soit  $c > 0$  comme dans la condition (3) de la définition des endomorphismes linéaires hyperboliques.

(1) et (2) Montrons l'égalité

$$E^s = \{x \in E : \forall n \in \mathbb{N}, \quad \|T^n(x)\| \leq c\lambda^n \|x\|\}, \quad (90)$$

dont l'inclusion directe est vérifiée par définition. Une démonstration analogue montre que

$$E^u = \{x \in E : \forall n \in \mathbb{N}, \quad \|x\| \leq c\lambda^n \|T^n(x)\|\}, \quad (91)$$

donc  $E^s$  et  $E^u$  sont uniquement déterminés. Par la continuité de la norme et de  $T$  donc de ses puissances, ceci montre aussi que  $E^s$  et  $E^u$  sont des intersections de fermés, donc sont fermés.

Soit  $x = x_s + x_u \in E$ , avec  $x_s \in E^s$  et  $x_u \in E^u$ , tel que  $\|T^n(x)\| \leq c\lambda^n \|x\|$  pour tout  $n \in \mathbb{N}$ . Pour tout  $n \in \mathbb{N}$ , nous avons  $T^n(x_u) \in E^u$  par la condition (2), donc par la condition (3), nous avons

$$\begin{aligned} \|x_u\| &= \|T^{-n}(T^n(x_u))\| \leq c\lambda^n \|T^n(x_u)\| = c\lambda^n \|T^n(x - x_s)\| \\ &\leq c\lambda^n (\|T^n(x)\| + \|T^n(x_s)\|) \leq c^2 \lambda^{2n} (\|x\| + \|x_s\|). \end{aligned}$$

Comme  $c^2 \lambda^{2n}$  tend vers 0 quand  $n$  tend vers  $+\infty$ , ceci implique que  $x_u = 0$  donc que  $x = x_s$  appartient à  $E^s$ . L'égalité (90) cherchée en découle.

(3) L'application  $\| \cdot \|' : E \rightarrow \mathbb{R}$  définie par

$$\forall x_s \in E^s, \forall x_u \in E^u, \quad \|x_s + x_u\|' = \max \left\{ \sum_{i=0}^{N-1} \kappa_s^{-i} \|T^i(x_s)\|, \sum_{i=0}^{N-1} \kappa_u^i \|T^{-i}(x_u)\| \right\}$$

pour  $N$  assez grand est une norme adaptée : il est facile de vérifier qu'il suffit de prendre  $N$  tel que  $\left(\frac{\kappa'_s}{\kappa_s}\right)^N, \left(\frac{\kappa'_u}{\kappa_u}\right)^N \leq \frac{1}{c}$ .  $\square$

Rappelons (voir par exemple les rappels de l'appendice 10.11 de ce chapitre) que le spectre  $\text{Sp}(T)$  de  $T$  est l'ensemble compact non vide des  $\lambda \in \mathbb{C}$  tels que, si  $E_{\mathbb{C}}$  est l'espace de Banach complexe complexifié de  $E$ , et si  $T_{\mathbb{C}}$  est l'endomorphisme linéaire continu de  $E_{\mathbb{C}}$  étendant  $T$ , alors  $T_{\mathbb{C}} - \lambda \text{id}$  soit inversible. Pour les parties 10.1, 10.2 et 10.3, le cas où  $E$  est de dimension finie suffit, auquel cas  $\text{Sp}(T)$  est l'ensemble des valeurs propres complexes de  $T$ .

**Proposition 10.2.** *L'endomorphisme linéaire continu  $T$  de  $E$  est hyperbolique si et seulement si son spectre  $\text{Sp}(T)$  ne rencontre pas le cercle unité  $\mathbb{S}_1 = \{z \in \mathbb{C} : |z| = 1\}$ .*

*Si  $0 < \kappa_s < 1 < \kappa_u$ , l'endomorphisme  $T$  est  $(\kappa_s, \kappa_u)$ -hyperbolique si et seulement si  $\text{Sp}(T)$  ne rencontre pas l'anneau  $\{z \in \mathbb{C} : \kappa_s \leq |z| \leq \kappa_u\}$ .*

**Démonstration.** Nous ne démontrons ici ce résultat que si  $E$  est de dimension finie, en renvoyant à l'appendice 10.11 de ce chapitre pour une démonstration dans le cas général. Tout compact ne rencontrant pas le cercle unité ne rencontre pas un anneau assez fin contenant le cercle unité dans son intérieur, donc la seconde assertion dans l'énoncé de la proposition 10.2 implique la première.

Supposons que  $T$  soit  $(\kappa_s, \kappa_u)$ -hyperbolique. Alors  $T$  n'a pas de valeur propre de module compris entre  $\kappa_s$  et  $\kappa_u$ , car si  $x = x_s + x_u$ , avec  $x_s \in E^s$  et  $x_u \in E^u$ , est un vecteur propre non nul de valeur propre  $\lambda$ , alors  $x_s$  ou  $x_u$  est un vecteur propre non nul de valeur propre  $\lambda$  par invariance des espaces stable et instable, et les relations  $T(x_s) = \lambda x_s$  ou  $T(x_u) = \lambda x_u$  contredisent l'assertion (3) de la définition des endomorphismes linéaires hyperboliques si  $\kappa_s \leq |\lambda| \leq \kappa_u$ .

Réciproquement, si  $T$  n'a pas de valeur propre de module compris entre  $\kappa_s$  et  $\kappa_u$ , alors par la théorie de la réduction<sup>202</sup> des endomorphismes réels en dimension finie, nous avons une décomposition en somme directe  $E = E^s \oplus E^u$  où  $E^s$  et  $E^u$  sont deux sous-espaces vectoriels réels invariants par  $T$ , tels que les valeurs propres complexes de  $T|_{E^s}$  et  $T|_{E^u}$  soient de module respectivement strictement inférieur à  $\kappa_s$  et strictement supérieur à  $\kappa_u$ . En particulier, l'endomorphisme  $T|_{E^u} : E^u \rightarrow E^u$  est inversible, car sans valeur propre nulle. La décomposition de Jordan<sup>203</sup> de  $T$  donne alors l'assertion (3) de la définition des

202. Si  $\chi_T = \det(X \text{id} - T) = \prod_{\lambda \in \text{Sp}(T)} (X - \lambda)$  est le polynôme caractéristique de  $T$ , si  $\chi_T = P_s P_u$  est la décomposition de  $\chi_T$  en produit de deux polynômes de coefficient dominant 1 telle que les racines de  $P_s = \prod_{\lambda \in \text{Sp}(T), |\lambda| < 1} (\lambda - X)$  (respectivement  $P_u = \prod_{\lambda \in \text{Sp}(T), |\lambda| > 1} (\lambda - X)$ ) soient les valeurs propres complexes de module strictement inférieur (respectivement supérieur) à 1, alors les polynômes  $P_s$  et  $P_u$  sont réels et premiers entre eux, donc le lemme de décomposition des noyaux dit que nous avons une décomposition en somme directe  $E = \ker P_s(T) \oplus \ker P_u(T)$ .

203. Un bloc de Jordan de taille  $d \in \mathbb{N} - \{0\}$  est une matrice  $J_d(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix} \in \mathcal{M}_d(\mathbb{C})$ ,

où  $\lambda \in \mathbb{C}$ . Notons  $I_d$  la matrice identité  $d \times d$ . Remarquons que  $J_d(\lambda) = \lambda I_d + J_d(0)$  avec  $J_d(0)$  triangulaire supérieure stricte. Un bloc de Jordan  $J_d(\lambda)$  est inversible si et seulement si  $\lambda \neq 0$ , et alors il existe une matrice triangulaire supérieure stricte  $P_\lambda$  telle que  $J_d(\lambda)^{-1} = \frac{1}{\lambda} I_d + P_\lambda$ .

La dynamique d'un bloc de Jordan, ainsi que de son inverse lorsqu'il est inversible, est donnée par les formules suivantes. Pour tout  $\mu \in \mathbb{C}$ , pour toute matrice  $P \in \mathcal{M}_d(\mathbb{C})$  triangulaire supérieure stricte,

endomorphismes linéaires hyperboliques avec

$$\kappa'_s \in ] \max\{|\lambda| : \lambda \in \text{Sp}(T), |\lambda| < 1\}, 1[$$

et

$$\kappa'_u \in ]1, \min\{|\lambda| : \lambda \in \text{Sp}(T), |\lambda| > 1\}[ . \quad \square$$

Le résultat suivant implique en particulier qu'une petite perturbation non linéaire (suffisamment régulière) d'un automorphisme linéaire hyperbolique de  $E$  lui est topologiquement conjugué (donc admet un comportement de dynamique topologique semblable, voir le chapitre 0). Nous rappelons que l'espace  $C^0_{ucb}(E, E)$  des applications uniformément continues bornées de  $E$  dans  $E$ , muni de la norme uniforme  $\|f\|_\infty = \sup_{x \in E} \|f(x)\|$ , est un espace de Banach.<sup>204</sup> Notons

$$\text{Lip}(f) = \sup_{x \neq y} \frac{\|f(x) - f(y)\|}{\|x - y\|}$$

pour tout  $n \geq d$ , puisque les matrices d'homothéties commutent avec toutes les matrices dans  $\mathcal{M}_d(\mathbb{C})$ , et puisque  $P$  est nilpotente d'ordre  $d$ , nous avons

$$(\mu I_d + P)^n = \mu^{n-d} \sum_{k=0}^{d-1} \mu^{d-k} P^k .$$

En particulier, pour la norme d'opérateurs sur  $\mathcal{M}_d(\mathbb{C})$ , pour tout  $\mu \in \mathbb{C}$  tel que  $|\mu| < 1$ , pour tout  $\kappa \in ]|\mu|, 1[$ , il existe  $c > 0$  tel que pour tout  $n \in \mathbb{N}$ , nous ayons

$$\|(\mu I_d + P)^n\| \leq c \kappa^n .$$

Par exemple, nous pouvons prendre

$$c = \max \left\{ \max_{1 \leq n \leq d-1} \|(\mu I_d + P)^n\|, \frac{\|\sum_{k=0}^{d-1} \mu^{d-k} P^k\|}{\kappa^d} \right\} .$$

Le *théorème de Jordan* réel dit que pour toute matrice  $M \in \mathcal{M}_N(\mathbb{R})$ , il existe une matrice inversible  $Q \in \text{GL}_N(\mathbb{C})$ , des entiers  $d_1, \dots, d_k \in \mathbb{N} - \{0\}$  tels que  $d_1 + \dots + d_k = N$  et des nombres complexes  $\lambda_1, \dots, \lambda_k \in \mathbb{C}$  (pas forcément deux à deux distincts) tels que les  $\lambda_i$  répétés  $d_i$  fois soient les valeurs propres complexes avec multiplicités de  $M$ , et tels que

$$M = Q \begin{pmatrix} J_{d_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J_{d_k}(\lambda_k) \end{pmatrix} Q^{-1} .$$

Nous pouvons supposer que  $0 \leq |\lambda_1| \leq \dots \leq |\lambda_\ell| < 1 < |\lambda_{\ell+1}| \leq \dots \leq |\lambda_k|$  pour un certain  $\ell \in \{0, \dots, k\}$ , quitte à permuter les blocs de Jordan. Avec les notations de la note de bas de page précédente 202, nous avons alors

$$E_s = \ker P_s(M) = Q \text{Vect}_{\mathbb{R}}(e_1, \dots, e_{d_1+d_2+\dots+d_\ell}) \quad \text{et} \quad E_u = \ker P_u(M) = Q \text{Vect}_{\mathbb{R}}(e_{d_1+d_2+\dots+d_{\ell+1}}, \dots, e_N) .$$

Comme

$$M^n = P \begin{pmatrix} J_{d_1}(\lambda_1)^n & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & J_{d_k}(\lambda_k)^n \end{pmatrix} P^{-1} .$$

pour tout  $n \in \mathbb{N}$ , l'affirmation finale de la démonstration de la proposition 10.2 en découle.

204. Rappelons qu'une application  $f : X \rightarrow Y$  entre deux espaces métriques est *uniformément continue* si

$$\forall \epsilon > 0, \exists \eta > 0, \forall x, y \in X, \quad \text{si } d(x, y) < \eta \quad \text{alors } d(f(x), f(y)) < \epsilon .$$

la *constante de Lipschitz* (globale) d'une application  $f : E \rightarrow E$ . Rappelons qu'une application  $f : E \rightarrow E$  est (globalement)  $\delta$ -lipschitzienne pour un  $\delta > 0$  si  $\text{Lip}(f) \leq \delta$ , qu'elle est (globalement) *lipschitzienne* s'il existe  $\delta > 0$  telle que  $f$  soit  $\delta$ -lipschitzienne, et qu'elle est (globalement) bilipschitzienne si elle est lipschitzienne, bijective, d'inverse lipschitzienne.

**Proposition 10.3. (Stabilité des automorphismes linéaires hyperboliques)** *Soit  $T$  un automorphisme linéaire continu hyperbolique d'un espace de Banach réel  $E$ , dont la norme est adaptée à  $T$ . Soit  $F = T + f$  où  $f : E \rightarrow E$  est une application bornée et  $\delta$ -lipschitzienne avec  $\delta < \delta_0 = \min\{1 - \text{ch}(T), \|T^{-1}\|^{-1}\}$ . Alors il existe un unique homéomorphisme  $H = \text{id} + h$  de  $E$  tel que  $h : E \rightarrow E$  soit uniformément continue et bornée, avec*

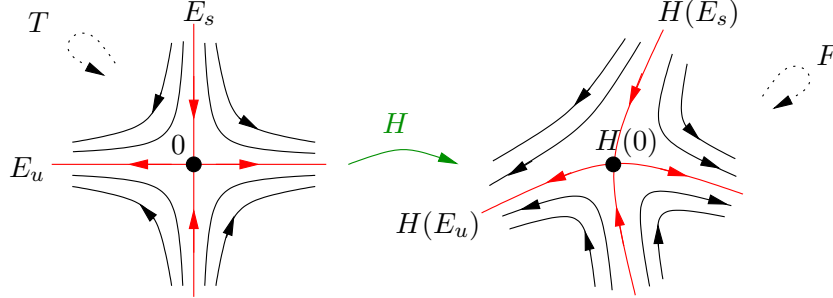
$$H \circ T \circ H^{-1} = F .$$

En particulier, les systèmes dynamiques topologiques  $(E, F)$  et  $(E, T)$  sont conjugués.

**Remarques.** (1) Puisque le vecteur nul de  $E$  est l'unique point fixe de  $T$  et par conjugaison, l'application  $F$  admet donc un unique point fixe : le point  $p = H(0)$ . Il résultera de la démonstration que

$$\|H(0)\| \leq \frac{\|f(0)\|}{\delta_0 - \delta} .$$

Les images par  $H$  de  $E^s$  et  $E^u$  sont invariantes par  $F$ . Privées de  $H(0)$ , elles sont respectivement contractées et dilatées par  $F$ .



(2) Il résultera de la démonstration que

$$\|h\|_\infty \leq \frac{1}{1 - \delta - \text{ch}(T)} \|f\|_\infty$$

et que  $h$  dépend continûment de  $f$  (pour la topologie de la norme uniforme).

(3) Il n'est pas toujours possible de trouver une telle application  $h$  qui soit lipschitzienne. Par exemple, prenons  $E = \mathbb{R}$ ,  $T : x \mapsto 2x$ ,  $f : x \mapsto -\epsilon \sin x$  avec  $\epsilon \in ]0, 2[$ , de sorte que

Le sous-espace vectoriel  $C_{ucb}^0(E, E)$  est un sous-espace fermé (donc complet) de l'espace de Banach  $C_b^0(E, E)$  des applications continues bornées de  $E$  dans  $E$ , muni de la norme uniforme. En effet, soit  $(f_n)_{n \in \mathbb{N}}$  une suite dans  $C_{ucb}^0(E, E)$  qui converge uniformément vers  $f \in C_b^0(E, E)$ . Soit  $\epsilon > 0$ . Fixons  $n$  assez grand tel que  $\|f - f_n\|_\infty \leq \frac{\epsilon}{3}$ . Puisque  $f_n$  est uniformément continue, il existe  $\delta > 0$  tel que pour tous les  $x, y \in E$ , si  $\|x - y\| \leq \delta$ , alors  $\|f_n(y) - f_n(x)\| \leq \frac{\epsilon}{3}$ . Donc par l'inégalité triangulaire, pour tous les  $x, y \in E$ , si  $\|x - y\| \leq \delta$ , alors

$$\|f(y) - f(x)\| \leq \|f(y) - f_n(y)\| + \|f_n(y) - f_n(x)\| + \|f_n(x) - f(x)\| \leq \epsilon .$$

$f$  est bornée et  $\epsilon$ -lipschitzienne. Alors l'homéomorphisme  $H$  donné par la Proposition 10.3 vérifie  $H(0) = 0$  (car 0 est l'unique point fixe de  $T$  et de  $F = T + f$ ). Donc  $h(0) = 0$ . Si nous notons  $x_n = 2^{-n}$  pour tout  $n \in \mathbb{N}$ , nous avons  $H(x_n) \neq 0$  et  $\lim_{n \rightarrow \infty} H(x_n) = 0$ . De plus,

$$\lim_{n \rightarrow \infty} \frac{H(x_{n+1})}{H(x_n)} = \lim_{n \rightarrow \infty} \frac{H(x_{n+1})}{H(2x_{n+1})} = \lim_{n \rightarrow \infty} \frac{H(x_{n+1})}{F(H(x_{n+1}))} = \frac{1}{F'(0)} = \frac{1}{2 - \epsilon} > \frac{1}{2}.$$

Puisque  $H = \text{id} + h$ , nous avons donc  $\lim_{n \rightarrow \infty} \frac{h(x_n) - h(0)}{x_n - 0} = \lim_{n \rightarrow \infty} \frac{H(x_n)}{x_n} - 1 = \infty$ .

**Démonstration.** Soit  $\delta < \min\{1 - \text{ch}(T), \|T^{-1}\|^{-1}\}$ . Pour des raisons d'argument de symétrie, soient  $F = T + f$  et  $G = T + g$  où  $f, g : E \rightarrow E$  sont deux applications bornées et  $\delta$ -lipschitziennes. Montrons que

- (i) il existe un unique élément  $h$  de  $C_{ucb}^0(E, E)$  tel que  $H = \text{id} + h$  vérifie  $F \circ H = H \circ G$ ,
- (ii)  $\|h\|_\infty \leq (1 - \delta - \text{ch}(T))^{-1} \|f - g\|_\infty$  et
- (iii)  $H$  est un homéomorphisme.

Le cas  $g = 0$  donne le résultat.

Nous commençons par démontrer un lemme de déformation d'automorphismes linéaires en homéomorphismes qui n'utilise pas l'hyperbolicité.

**Lemme 10.4.** Soient  $T' \in \mathcal{GL}(E)$  et  $g' : E \rightarrow E$  une application  $\delta'$ -lipschitzienne avec  $\delta' < \|T'^{-1}\|^{-1}$ . Alors l'application  $G' = T' + g'$  est un homéomorphisme bilipschitzien.

**Démonstration.** Comme  $G' = T' \circ (\text{id} + T'^{-1} \circ g')$  et  $\text{Lip}(T'^{-1} \circ g') \leq \|T'^{-1}\| \text{Lip}(g')$ , il suffit de montrer ce lemme avec  $T' = \text{id}$  et  $\delta' < 1$ , ce que nous supposons désormais.

Soient  $y \in E$  et  $\varphi : x \mapsto y - g'(x)$ . Puisque  $\delta' < 1$ , l'application  $\varphi$  est contractante sur l'espace métrique complet  $E$ . Elle admet donc<sup>205</sup> un unique point fixe  $x$ . Celui-ci vérifie  $G'(x) = x + g'(x) = y$ . Donc  $G'$  est surjective.

L'application  $G'$  est clairement  $(1 + \delta')$ -lipschitzienne. Pour tous les  $x, y \in E$ , nous avons

$$\|G'(x) - G'(y)\| \geq \|x - y\| - \|g'(x) - g'(y)\| \geq (1 - \delta') \|x - y\|.$$

Donc  $G'$  est injective, et  $G'^{-1}$  est  $(1 - \delta')^{-1}$ -lipschitzienne.  $\square$

(i) Pour toute application  $\gamma : E \rightarrow E$ , notons  $\gamma_s : E \rightarrow E^s$  et  $\gamma_u : E \rightarrow E^u$  les applications telles que  $\gamma = \gamma_s + \gamma_u$ . Notons  $S$  et  $U$  les restrictions de  $T$  à  $E^s$  et  $E^u$  respectivement. En utilisant les relations  $F = T + f$ ,  $G = T + g$  et  $H = \text{id} + h$ , l'égalité  $F \circ H = H \circ G$  s'écrit  $f \circ H + T \circ h = g + h \circ G$ . En projetant sur  $E^s$  et  $E^u$ , cette égalité est équivalente au système

$$\begin{cases} f_s \circ H + S \circ h_s &= g_s + h_s \circ G \\ f_u \circ H + U \circ h_u &= g_u + h_u \circ G, \end{cases}$$

la première équation étant une égalité entre applications de  $E$  dans  $E^s$ , la seconde entre applications de  $E$  dans  $E^u$ . Ce système est équivalent à l'égalité  $h = \tilde{h}$  entre applications de  $E$  dans  $E$ , où

$$\begin{cases} \tilde{h}_s &= (f_s \circ H + S \circ h_s - g_s) \circ G^{-1} \\ \tilde{h}_u &= U^{-1} \circ (h_u \circ G + g_u - f_u \circ H). \end{cases}$$

<sup>205</sup>. Voir les notes de bas de page (161) et (162).

Notons que par le lemme 10.4, puisque  $\delta < \|T^{-1}\|^{-1}$ , l'application  $G$  est un homéomorphisme bilipschitzien de  $E$ .

Montrons maintenant que l'application  $\varphi : C_{ucb}^0(E, E) \rightarrow C_{ucb}^0(E, E)$  définie par  $h \mapsto \tilde{h}$  est  $(\delta + \text{ch}(T))$ -lipschitzienne. Comme la norme de  $E$  est adaptée à  $T$ , pour évaluer  $\|\tilde{h} - \tilde{h}'\|_\infty$  avec  $h, h'$  dans  $C_{ucb}^0(E, E)$ , il suffit d'évaluer  $\|\tilde{h}_s - \tilde{h}'_s\|_\infty$  et  $\|\tilde{h}_u - \tilde{h}'_u\|_\infty$ . Posons  $H' = \text{id} + h'$ , de sorte que  $H - H' = h - h'$ . Puisque les applications  $f_s$  et  $f_u$  sont  $\delta$ -lipschitziennes et par la formule (89), nous avons

$$\begin{aligned} \|\tilde{h}_s - \tilde{h}'_s\|_\infty &\leq \|f_s \circ H - f_s \circ H'\|_\infty + \|S\| \|h_s - h'_s\|_\infty \leq (\text{Lip}(f_s) + \|S\|) \|h - h'\|_\infty \\ &\leq (\delta + \text{ch}(T)) \|h - h'\|_\infty. \end{aligned}$$

Nous avons  $\|U^{-1}\| < 1$ , puisque la norme de  $E$  est adaptée à  $T$ . Donc

$$\begin{aligned} \|\tilde{h}_u - \tilde{h}'_u\|_\infty &\leq \|U^{-1}\| \|h_u \circ G - h'_u \circ G\|_\infty + \|U^{-1}\| \|f_u \circ H - f_u \circ H'\|_\infty \\ &\leq (\|U^{-1}\| + \text{Lip}(f_u)) \|h - h'\|_\infty \leq (\delta + \text{ch}(T)) \|h - h'\|_\infty. \end{aligned}$$

Puisque  $\delta + \text{ch}(T) < 1$ , l'existence et l'unicité d'une solution  $h : E \rightarrow E$  de l'équation  $h = \tilde{h}$  résulte alors du théorème 8.10 (dans la note de bas de page (162)) du point fixe des applications contractantes, appliqué à l'application  $(\delta + \text{ch}(T))$ -lipschitzienne  $\varphi$ .

(ii) La majoration de  $\|h\|_\infty$  s'obtient à l'aide de la dernière assertion du théorème 8.10, qui dit que

$$d(h, x_0) \leq \frac{1}{1 - (\delta + \text{ch}(T))} d(\varphi(x_0), x_0)$$

pour tout  $x_0 \in C_{ucb}^0(E, E)$ , en prenant  $x_0 = 0$ . En effet, puisque la norme est adaptée, nous avons

$$d(\varphi(0), 0) = \|\tilde{0}\|_\infty \leq \max\{\|(f_s - g_s) \circ G^{-1}\|_\infty, \|U^{-1}\| \|g_u - f_u\|_\infty\} \leq \|g - f\|_\infty.$$

(iii) Montrons pour conclure que  $H$  est un homéomorphisme. En échangeant les rôles de  $f$  et  $g$ , par l'assertion (i), nous obtenons un élément  $h' \in C_{ucb}^0(E, E)$  tel que  $H' = \text{id} + h'$  vérifie  $H' \circ F = G \circ H'$ . Nous avons donc  $(H \circ H') \circ F = F \circ (H \circ H')$ . Comme  $H \circ H' - \text{id} = h' + h \circ H'$  appartient à  $C_{ucb}^0(E, E)$ , par la propriété d'unicité dans l'assertion (i), nous avons  $H \circ H' = \text{id}$ . De même, nous avons  $H' \circ H = \text{id}$ , et  $H$  est un homéomorphisme.  $\square$

Nous aurons besoin dans le théorème 10.20 d'un meilleur contrôle sur l'unique point fixe  $p$  de  $F$  (voir la remarque (1) ci-dessus).

**Proposition 10.5.** *Avec les hypothèses de la proposition 10.3, l'unique point fixe  $p$  de  $F$  vérifie, pour tout  $v \in E$ , l'inégalité*

$$\|p - v\| \leq \frac{\|F(v) - v\|}{\delta_0 - \delta}.$$

**Démonstration.** Avec les notations de la démonstration de la proposition 10.3, identifions  $E^s \times E^u$  avec  $E$  par  $(x_s, x_u) \mapsto x_s + x_u$  et notons  $\bar{F} : E \rightarrow E$  l'application définie par

$$\bar{F}(x_s, x_u) = (F_s(x_s, x_u), x_u + U^{-1}(x_u - F_u(x_s, x_u))).$$

Alors  $\bar{F}$  et  $F$  ont les mêmes points fixes. Puisque  $F = T + f$ , nous avons

$$\bar{F}(x_s, x_u) = (Sx_s + f_s(x_s, x_u), U^{-1}(x_u - f_u(x_s, x_u))).$$

Puisque les applications  $f_s$  et  $f_u$  sont  $\delta$ -lipschitziennes et  $\|U^{-1}\| \leq 1$ , nous avons donc

$$\|\bar{F}_s(x_s, x_u) - \bar{F}_s(x'_s, x'_u)\| \leq (\|S\| + \delta)\|(x_s, x_u) - (x'_s, x'_u)\|$$

et

$$\begin{aligned} \|\bar{F}_u(x_s, x_u) - \bar{F}_u(x'_s, x'_u)\| &\leq \|U^{-1}\|(1 + \delta)\|(x_s, x_u) - (x'_s, x'_u)\| \\ &\leq (\|U^{-1}\| + \delta)\|(x_s, x_u) - (x'_s, x'_u)\|. \end{aligned}$$

Donc par la formule (89), l'application  $\bar{F}$  est  $(\delta + \text{ch}(T))$ -lipschitzienne.

Remarquons que, puisque la norme est adaptée et  $\|U^{-1}\| \leq 1$ , nous avons

$$\|\bar{F}(x_s, x_u) - (x_s, x_u)\| = \|F_s(x_s, x_u) - x_s, U^{-1}(x_u - F_u(x_s, x_u))\| \leq \|F(x_s, x_u) - (x_s, x_u)\|.$$

Le fait que l'unique point fixe de  $F$ , qui est l'unique point fixe de  $\bar{F}$ , vérifie la proposition 10.5 découle alors de la dernière assertion du théorème 8.10, car  $1 - (\delta + \text{ch}(T)) \geq \delta_0 - \delta$  par la définition de  $\delta_0$ .  $\square$

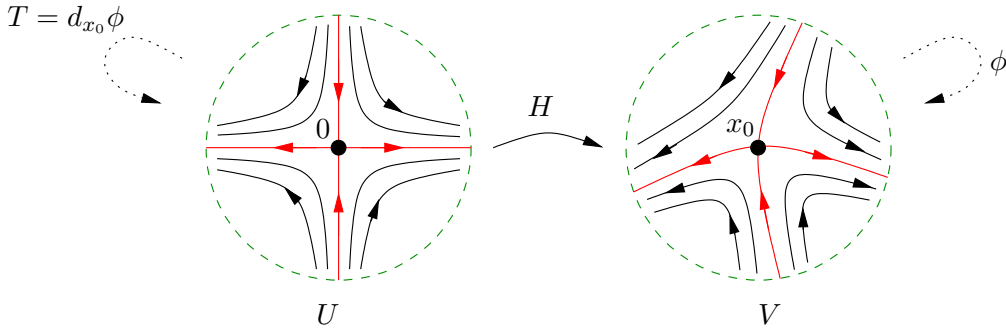
## 10.2 Le théorème de Grobman-Hartman

Le théorème suivant affirme que, dans certains cas, un difféomorphisme est localement topologiquement conjugué au voisinage d'un point fixe à sa différentielle en ce point fixe. En particulier, la dynamique de ce difféomorphisme sera, dans ce voisinage, "la même" que celle de sa différentielle. Voir [KH, §6.3] pour une autre démonstration.

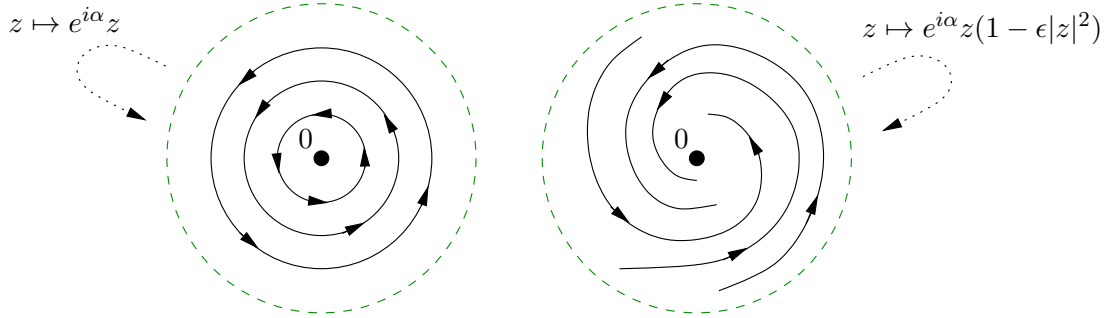
**Théorème 10.6. (Théorème de Grobman-Hartman)** *Soient  $N \in \mathbb{N} - \{0\}$ ,  $\Omega$  un ouvert de  $\mathbb{R}^N$ ,  $\phi : \Omega \rightarrow \mathbb{R}^N$  un  $C^1$ -difféomorphisme local et  $x_0$  un point fixe hyperbolique de  $\phi$  (c'est-à-dire  $x_0$  est un point de  $\Omega$  tel que  $\phi(x_0) = x_0$  et tel que la différentielle  $T = d_{x_0}\phi$  de  $\phi$  en  $x_0$  soit un automorphisme linéaire hyperbolique de  $\mathbb{R}^N$ ).*

*Alors il existe des voisinages ouverts  $U$  de  $0$  dans  $\mathbb{R}^N$  et  $V$  de  $x_0$  dans  $\Omega$ , et un homéomorphisme  $H : U \rightarrow V$  tel que  $H(0) = x_0$  et, pour tout  $x \in U$ , si  $T(x) \in U$ , alors*

$$H \circ T(x) = \phi \circ H(x).$$



**Remarque.** Si le point fixe  $x_0$  n'est pas hyperbolique, ceci n'est plus vrai en général. Par exemple, pour tout  $\epsilon \in ]0, 1]$ , nous pouvons considérer l'application  $\phi : \mathbb{C} \rightarrow \mathbb{C}$  définie par  $z \mapsto e^{i\alpha}z(1 - \epsilon|z|^2)$ , qui n'est topologiquement conjuguée à la rotation  $z \mapsto e^{i\alpha}z$  dans aucun voisinage de  $0$  (regarder quels sont les ensembles  $\omega$ -limites).



**Démonstration.** Soit  $\alpha : \mathbb{R}^N \rightarrow \mathbb{R}$  une fonction plateau de classe  $C^\infty$  telle que  $\alpha(x) = 1$  si  $\|x\| \leq 1$  et  $\alpha(x) = 0$  si  $\|x\| \geq 2$ . Pour tout  $\epsilon > 0$ , soit  $\alpha_\epsilon(x) = \alpha(\frac{x}{\epsilon})$ .

Nous pouvons supposer que  $x_0 = 0$ . Écrivons alors  $\phi = T + f$  où  $f : \Omega \rightarrow \mathbb{R}^N$  est de classe  $C^1$  et vérifie  $f(0) = 0$  et  $d_0 f = 0$ . Posons  $F_\epsilon = T + f_\epsilon$  avec  $f_\epsilon = \alpha_\epsilon f$ . Prenons  $\epsilon$  assez petit. L'application  $F_\epsilon : \mathbb{R}^N \rightarrow \mathbb{R}^N$  coïncide avec  $\phi$  sur  $B(0, \epsilon)$  et avec  $T$  en dehors de  $B(0, 2\epsilon)$ . L'application  $f_\epsilon$  est bornée et lipschitzienne, avec (par le théorème des accroissements finis)

$$\text{Lip}(f_\epsilon) = \sup_{x \in \mathbb{R}^N} \|d_x f_\epsilon\|.$$

Or nous avons

$$d_x f_\epsilon = \alpha\left(\frac{x}{\epsilon}\right) d_x f + \frac{f(x)}{\epsilon} d_x \alpha.$$

Notons  $M = \max \{ \|\alpha\|_\infty, \|x \mapsto \|d_x \alpha\|_\infty \}$ . Nous avons alors

$$\text{Lip}(f_\epsilon) \leq M \left( \sup_{x \in B(0, 2\epsilon)} \|d_x f\| + \sup_{x \in B(0, 2\epsilon)} \frac{\|f(x)\|}{\epsilon} \right) \leq 3M \sup_{x \in B(0, 2\epsilon)} \|d_x f\|$$

(encore par le théorème des accroissements finis, car  $f(0) = 0$ ). Puisque  $d_0 f = 0$ , la constante de Lipschitz  $\text{Lip}(f_\epsilon)$  tend donc vers 0 quand  $\epsilon \rightarrow 0$ , et nous pouvons appliquer la proposition 10.3 à  $F_\epsilon$  pour  $\epsilon > 0$  fixé assez petit. Il existe donc un homéomorphisme  $H$  de  $\mathbb{R}^N$  tel que  $H \circ T = F_\epsilon \circ H$ . Puisque  $T$  admet 0 comme seul point fixe, et puisque  $F_\epsilon(0) = 0$ , nous avons par conjugaison  $H(0) = 0$ . Nous prenons alors  $U = B(0, \epsilon)$  et  $V = H(B(0, \epsilon))$  et nous remplaçons  $H$  par sa restriction à  $U$  afin d'obtenir le résultat.  $\square$

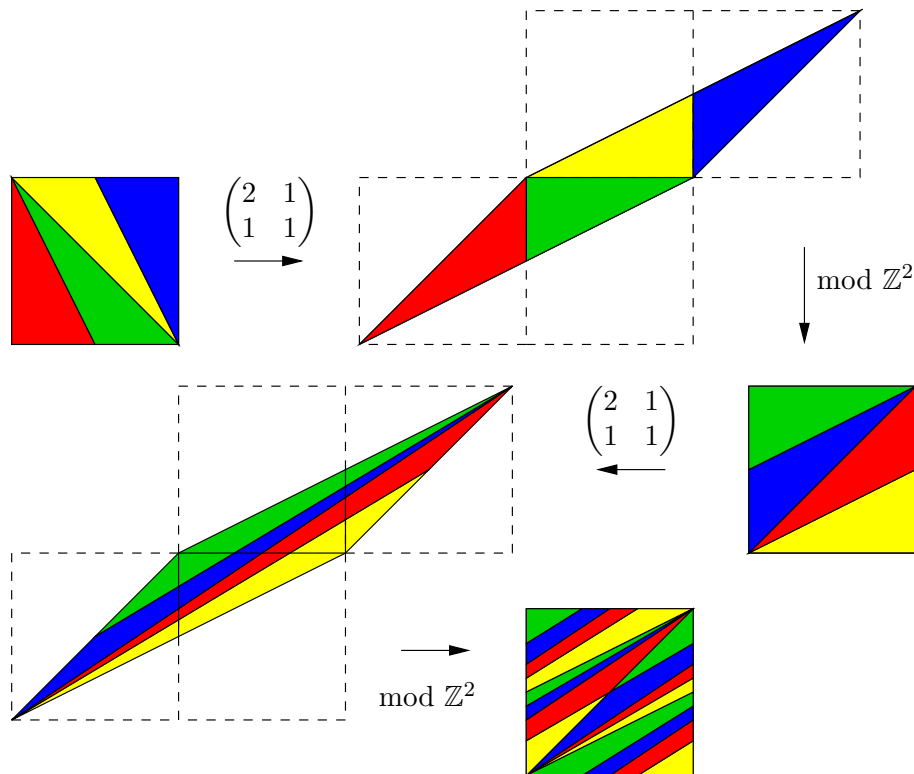
### 10.3 Stabilité structurelle des automorphismes hyperboliques du tore

Avant de donner la définition générale d'un système dynamique (uniformément) hyperbolique dans la partie 10.4, donnons un résultat de stabilité globale de dynamique chaotique. C'est un cas particulier d'un résultat plus général que nous énoncerons ultérieurement, mais dont la démonstration est plus simple.

Soient  $N \in \mathbb{N} - \{0\}$  et  $M \in \mathcal{M}_N(\mathbb{Z})$  une matrice carrée à coefficients entiers de déterminant non nul de taille  $N$ . Nous avons vu dans les chapitres précédents que la dynamique de la transformation  $\phi_M : \mathbb{T}^N \rightarrow \mathbb{T}^N$  (définie dans la partie 4) peut être très chaotique : mélange, densité des orbites périodiques, mélange multiple, mélange exponentiel, ... Le théorème suivant affirme que, lorsque  $M$  est hyperbolique, cette dynamique chaotique n'est pas modifiée par une petite perturbation de  $\phi_M$ .



L'application du chat d'Arnold, qui est un des exemples le plus étudié de système dynamique hyperbolique, correspond au cas  $N = 2$  et  $M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ . L'une des raisons est sans doute que les valeurs propres complexes de  $M$  sont le carré du nombre d'or  $\varphi^2 = \frac{3+\sqrt{5}}{2} > 1$  et son conjugué de Galois  $\frac{3-\sqrt{5}}{2} < 1$  (qui ne sont donc pas de module 1). La matrice  $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ , dont le carré est la matrice précédente, conviendrait aussi bien (voire mieux du point de vue des valeurs propres complexes!), mais pour des raisons historiques, c'est la précédente qui fait l'objet de nombreux documents.

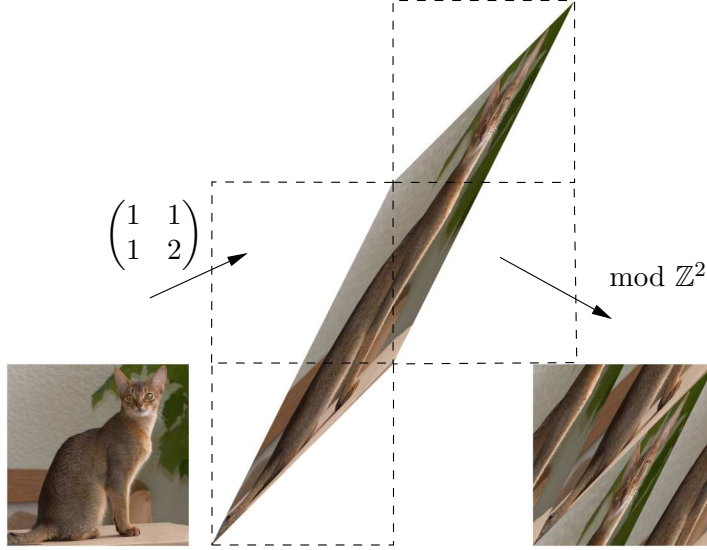


La transformation linéaire de matrice  $M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$  envoie le carré unité  $[0, 1]^2$  sur le parallélogramme de sommets  $(0, 0)$ ,  $(2, 1)$ ,  $(3, 2)$  et  $(1, 1)$ , comme il est facile de vérifier en regardant l'image des sommets du carré. Pour comprendre l'application  $\phi_M$  induite par  $M$  sur le tore  $\mathbb{T}^N$ , il suffit alors de réduire modulo  $\mathbb{Z}^2$ , voir le dessin ci-dessus pour une visualisation de  $\phi_M^2$  : nous avons vraiment l'impression que  $M$  mélange les couleurs, ce qui illustre le théorème 4.10.

Comme dit ci-dessus, ce sont les propriétés fortes de contraction et dilatation, couplées avec de la récurrence, qui provoquent ce phénomène. Les deux droites propres de  $M$  sont  $E^u = \mathbb{R}(\frac{1+\sqrt{5}}{2}, 1)$ , dilatée par le facteur  $\varphi^2 = \frac{3+\sqrt{5}}{2} > 1$  et  $E^s = \mathbb{R}(\frac{1-\sqrt{5}}{2}, 1)$ , contractée par le facteur  $\frac{3-\sqrt{5}}{2} < 1$ . Nous reviendrons sur une description générale de la décomposition dynamique  $\mathbb{R}^2 = E^s \oplus E^u$  dans la partie 10.4.

En hommage à Arnold, qui en 1967 introduisit une tête stylisée de chat dans le carré unité  $[0, 1]^2$ , et dessina son image par  $\phi_M$ , ce qui donna son nom à l'application  $\phi_M$ , voici

une image de transformation de chat, extraite de Wikipédia (pour une matrice légèrement différente de  $M$ , mais conjuguée, donc ayant les mêmes valeurs propres complexes).



Nous aurons besoin du rappel de topologie algébrique élémentaire suivant, pour lequel nous renvoyons par exemple à [Pau4, Pau8], ainsi que [God1, Spa, Hat], définissant, pour toute application continue  $\phi$  du tore  $\mathbb{T}^N$  dans lui-même, d'une part ses relèvements  $\tilde{\phi}$  au revêtement universel  $\mathbb{R}^N$  de  $\mathbb{T}^N$ , et d'autre part son action  $\phi_*$  sur le groupe fondamental  $\mathbb{Z}^N$  de  $\mathbb{T}^N$ .

**Proposition 10.7.** Soient  $\phi : \mathbb{T}^N \rightarrow \mathbb{T}^N$  une application continue et  $\rho : \mathbb{R}^N \rightarrow \mathbb{T}^N$  la projection canonique. Appelons un relèvement de  $\phi$  par  $\rho$  toute application continue  $\tilde{\phi} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  telle que  $\rho \circ \tilde{\phi} = \phi \circ \rho$ .

(1) Pour tout les points  $x_0$  et  $x'_0$  de  $\mathbb{R}^N$  tels que  $\rho(x'_0) = \phi(\rho(x_0))$ , il existe un et un seul relèvement  $\tilde{\phi}$  de  $\phi$  tel que  $\tilde{\phi}(x_0) = x'_0$ .<sup>206</sup> De plus, si  $\phi$  est un homéomorphisme de  $\mathbb{T}^N$ , alors  $\tilde{\phi}$  est un homéomorphisme de  $\mathbb{R}^N$ .<sup>207</sup>

(2) Il existe une matrice  $\phi_* \in \mathcal{M}_N(\mathbb{Z})$  à coefficients entiers telle que

$$\forall x \in \mathbb{R}^N, \forall z \in \mathbb{Z}^N, \quad \tilde{\phi}(x + z) = \tilde{\phi}(x) + \phi_*(z). \quad ^{208}$$

206. Ceci découle du fait que la projection canonique  $\rho$  est un revêtement, et que  $\mathbb{R}^N$  est connexe par arcs et simplement connexe, voir les références ci-dessus, dont [Pau8, Coro. 2.26]. Notons que le point  $x_0 \in \mathbb{R}^N$  étant fixé (et nous prendrons souvent 0), le relèvement  $\tilde{\phi}$  dépend du choix de  $x'_0$ . Mais si  $x''_0$  est un autre point de  $\mathbb{R}^N$  tel que  $\rho(x''_0) = \phi(\rho(x_0))$  et si  $\tilde{\phi}^*$  est le relèvement associé à ce nouveau choix, alors il existe  $m \in \mathbb{Z}^N$  tel que  $x''_0 = x'_0 + m$ , et les deux relèvements  $\tilde{\phi}^*$  et  $\tilde{\phi}$  diffèrent seulement de la constante additive  $m$  : nous avons  $\tilde{\phi}^* = \tilde{\phi} + m$  par unicité.

207. En notant  $\tilde{\phi}^{-1}$  l'unique relèvement de l'application continue  $\phi^{-1}$  tel que  $\tilde{\phi}^{-1}(x'_0) = x_0$ , ce qui est possible car  $\rho(x_0) = \phi^{-1}(\rho(x'_0)) = \phi^{-1}(\rho(x'_0))$ , nous avons  $\tilde{\phi}^{-1} \circ \tilde{\phi} = \text{id}_{\mathbb{R}^N}$  par unicité du relèvement de l'identité fixant  $x_0$ , et de même  $\tilde{\phi} \circ \tilde{\phi}^{-1} = \text{id}_{\mathbb{R}^N}$ .

208. Soient  $z, z'$  dans  $\mathbb{Z}^N$ . L'application  $x \mapsto \tilde{\phi}(x + z) - \tilde{\phi}(x)$  est une application continue sur l'espace topologique connexe  $\mathbb{R}^N$  à valeurs dans l'espace topologique discret  $\mathbb{Z}^N$ . Elle est donc égale à une constante

De plus, si  $\psi : \mathbb{T}^N \rightarrow \mathbb{T}^N$  est une autre application continue, alors  $(\phi \circ \psi)_* = \phi_* \circ \psi_*$ .  
 En particulier, si  $\phi$  est un homéomorphisme de  $\mathbb{T}^N$ , alors pour tout  $n \in \mathbb{Z}$ , nous avons  $(\phi^n)_* = (\phi_*)^n$ .

(3) La matrice  $\phi_*$  est l'unique matrice  $M \in \mathcal{M}_N(\mathbb{Z})$  telle que l'application  $\tilde{\phi} - M$  soit bornée sur  $\mathbb{R}^N$ . [209](#)

(4) Si le tore  $\mathbb{T}^N$  est muni de la distance quotient  $d(\theta, \theta') = \inf_{\substack{x \in \mathbb{R}^N, \mathcal{P}(x) = \theta \\ x' \in \mathbb{R}^N, \mathcal{P}(x') = \theta'}} \|x - x'\|$ , si

$\phi, \psi : \mathbb{T}^N \rightarrow \mathbb{T}^N$  sont deux applications continues telles que  $\sup_{\theta \in \mathbb{T}^N} d(\phi(\theta), \psi(\theta)) < \frac{1}{2}$ , alors  $\phi_* = \psi_*$ . [210](#)  $\square$

La différentiabilité des applications de la variété différentielle  $\mathbb{T}^N$  dans elle-même (voir un cours de géométrie différentielle, par exemple [[Laf, Pau2](#)]) s'interprète facilement. Notons  $C^1(\mathbb{T}^N, \mathbb{T}^N)$  l'ensemble des applications  $f : \mathbb{T}^N \rightarrow \mathbb{T}^N$  de classe  $C^1$ , c'est-à-dire telles que  $\tilde{f} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  est de classe  $C^1$ , ce qui ne dépend pas du choix de relèvement  $\tilde{f}$  de  $f$ , puisque deux relèvements diffèrent d'une constante, comme vu dans la proposition [10.7 \(1\)](#) (et sa note de bas de page [206](#)).

Pour tout point  $\theta$  dans  $\mathbb{T}^N$ , nous notons alors  $T_\theta f = d_x \tilde{f} \in \mathcal{L}(\mathbb{R}^N)$  où  $x \in \mathbb{R}^N$  est n'importe quel point de  $\mathbb{R}^N$  tel que  $\theta = \mathcal{P}(x)$ . Par les propriétés rappelées ci-dessus, cette application linéaire  $T_\theta f$  ne dépend ni du choix du relèvement  $\tilde{f}$  de  $f$  ni du choix du point  $x$  dans la fibre  $\mathcal{P}^{-1}(\theta)$  au-dessus de  $\theta$ . C'est, par définition, la *différentielle* de  $f$  en  $\theta$ .

En utilisant la norme d'opérateur  $\| \cdot \|$  sur l'algèbre  $\mathcal{L}(\mathbb{R}^N)$ , nous munissons l'ensemble  $C^1(\mathbb{T}^N, \mathbb{T}^N)$  de la distance  $d$  donnée par

$$d(f, g) = \sup_{\theta \in \mathbb{T}^N} ( d(f(\theta), g(\theta)) + \|T_\theta f - T_\theta g\| ).$$

La topologie induite par cette distance est appelée la *topologie*  $C^1$  sur  $C^1(\mathbb{T}^N, \mathbb{T}^N)$  (voir la généralisation de cette notion à toutes les variétés, avant l'énoncé du corollaire [10.16](#)).

**Exemple.** Soit  $M \in \mathcal{M}_N(\mathbb{Z})$  une matrice à coefficients entiers, de déterminant non nul. Puisque la différentielle d'une application linéaire en tout point est elle-même, puisque  $\phi_M(0) = 0$  et en prenant  $x_0 = x'_0 = 0$  dans la proposition [10.7 \(1\)](#), nous avons  $\phi_M = M$  et  $T_\theta \phi_M = M$  pour tout  $\theta \in \mathbb{T}^N$ . Par l'unicité dans la proposition [10.7 \(3\)](#), nous avons  $(\phi_M)_* = M$ .

Le résultat suivant est le théorème de stabilité structurelle d'Anosov dans le cas particulier des difféomorphismes d'Anosov linéaires des tores. Nous renvoyons par exemple à [[KH, 18.2](#)] pour une version générale.

$\phi_*(z) \in \mathbb{Z}^N$ . Pour n'importe quel  $x \in \mathbb{R}^N$ , nous avons

$$\phi_*(z) + \phi_*(z') = (\tilde{\phi}(x + z' + z) - \tilde{\phi}(x + z')) + (\tilde{\phi}(x + z') - \tilde{\phi}(x)) = \phi_*(z + z').$$

Donc  $\phi_*$  est un morphisme de groupes de  $\mathbb{Z}^N$  dans lui-même, et est donc donnée par une matrice à coefficients entiers.

209. Ceci découle du fait que  $\tilde{\phi} - \phi_*$  est une application continue et  $\mathbb{Z}^N$ -périodique par l'assertion [\(2\)](#), donc bornée, et du fait que toute application linéaire bornée est nulle.

210. Notons  $\tilde{\phi}, \tilde{\psi} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  des relèvements de  $\phi$  et  $\psi$ . Par l'hypothèse de l'assertion [\(4\)](#), l'application continue  $\tilde{\phi} - \tilde{\psi}$  prend ses valeurs dans  $\bigcup_{z \in \mathbb{Z}^N} \overset{\circ}{B}(z, \frac{1}{2})$ . Comme  $\mathbb{R}^N$  est connexe et puisque ces boules ouvertes sont disjointes, l'image de l'application  $\tilde{\phi} - \tilde{\psi}$  est contenue dans une de ces boules, donc est bornée. Par conséquent, par l'assertion [\(3\)](#), l'application linéaire  $\phi_* - \psi_*$  est bornée, donc nulle.

**Théorème 10.8. (Anosov)** Soit  $M \in \mathcal{M}_N(\mathbb{Z})$  une matrice à coefficients entiers, de déterminant non nul et hyperbolique. Alors il existe un voisinage  $\mathcal{U}$  de  $\phi_M$  dans  $C^1(\mathbb{T}^N, \mathbb{T}^N)$  tel que toute application  $\phi$  dans  $\mathcal{U}$  est topologiquement conjuguée à  $\phi_M$  : il existe un et un seul homéomorphisme  $h$  de  $\mathbb{T}^N$  tel que  $h \circ \phi = \phi_M \circ h$  et  $h_* = \text{id}$ .

**Démonstration.** D'après la proposition 10.7 (4), si nous choisissons  $\mathcal{U}$  suffisamment petit, tous les éléments  $\phi$  de  $\mathcal{U}$  vérifient

$$\phi_* = (\phi_M)_* = M .$$

D'après la proposition 10.7 (3), les relèvements  $\tilde{\phi}$  de  $\phi$  s'écrivent  $\tilde{\phi} = M + \psi$  avec  $\psi$  une application bornée sur  $\mathbb{R}^N$ . Quitte à réduire  $\mathcal{U}$ , nous pouvons supposer que la constante de Lipschitz

$$\text{Lip}(\psi) = \sup_{\theta \in \mathbb{T}^N} \|T_\theta \psi\| = \sup_{\theta \in \mathbb{T}^N} \|T_\theta \phi - T_\theta \phi_M\|$$

est aussi petite que l'on veut.

La proposition 10.3 nous donne alors un homéomorphisme  $H$  de  $\mathbb{R}^N$  tel que  $H \circ \tilde{\phi} = M \circ H$  et tel que l'application  $H - \text{id}$  est bornée.

**Lemme 10.9.** Soit  $x$  dans  $\mathbb{R}^N$ . Le point  $H(x)$  est l'unique point  $y$  de  $\mathbb{R}^N$  tel que

$$\sup_{n \in \mathbb{Z}} \|M^n y - \tilde{\phi}^n(x)\| < \infty .$$

**Démonstration.** Tout d'abord, le point  $H(x)$  vérifie bien cette condition, car nous avons  $M^n H(x) - \tilde{\phi}^n(x) = H(\tilde{\phi}^n(x)) - \tilde{\phi}^n(x)$  et l'application  $H - \text{id}$  est bornée.

Réciproquement, si un point  $y = H(x) + v$  vérifie cette condition, alors la suite  $(M^n v)_{n \in \mathbb{Z}}$  est bornée. Comme  $M$  est hyperbolique, nous avons forcément  $v = 0$ .  $\square$

Pour terminer la démonstration du théorème d'Anosov 10.8, montrons que la conjugaison  $H$  entre  $M$  et  $\tilde{\phi}$  passe au quotient en une conjugaison  $h$  entre  $\phi_M$  et  $\phi$  : montrons qu'il existe une application continue  $h : \mathbb{T}^N \rightarrow \mathbb{T}^N$  telle que  $\mathcal{r} \circ H = h \circ \mathcal{r}$ . Pour cela, il suffit de montrer que

$$\forall z \in \mathbb{Z}^N, \forall x \in \mathbb{R}^N, \quad H(x + z) = H(x) + z .$$

Mais cette égalité résulte du lemme ci-dessus, car la suite indexée par  $n \in \mathbb{Z}$  des

$$M^n(H(x) + z) - \tilde{\phi}^n(x + z) = (M^n H(x) + M^n z) - (\tilde{\phi}^n(x) + \phi_*^n(z)) = M^n H(x) - \tilde{\phi}^n(x)$$

est bornée.

L'application  $h$  vérifie  $h \circ \phi = \phi_M \circ h$  par passage au quotient. Il reste à voir que  $h$  est un homéomorphisme. Or l'inverse  $H' = H^{-1}$  vérifie aussi les égalités

$$\forall z \in \mathbb{Z}^N, \forall x \in \mathbb{R}^N, \quad H'(x + z) = H'(x) + z .$$

Il existe donc aussi une application continue  $h'$  telle que  $\mathcal{r} \circ H' = h' \circ \mathcal{r}$ . Par construction,  $h'$  est l'inverse de  $h$ .  $\square$

## 10.4 Ensembles hyperboliques de difféomorphismes

Nous renvoyons par exemple à [Laf, Pau2] pour les prérequis sur les variétés différentielles, les fibrés tangents et les champs de vecteurs.

### 10.4.1 Fibrés vectoriels normés

Nous commençons cette partie par des “rappels” sur les fibrés vectoriel normés, que le lecteur savant (ou pressé) peut omettre pour commencer après la fin de la démonstration de la proposition 10.10, il n’y en a pas vraiment besoin pour “comprendre” la suite.

Si  $K$  est un espace topologique, un *fibré vectoriel* (réel) sur  $K$  est un espace topologique  $E$ , appelé l’*espace total*, muni d’une application continue surjective  $\pi : E \rightarrow K$ , appelée la *projection*, et, pour tout  $x \in K$ , d’une structure d’espace vectoriel réel sur la *fibres*  $E_x = \pi^{-1}(x)$ , vérifiant la propriété de *trivialité locale* suivante : pour tout  $x \in K$ , il existe un voisinage ouvert  $U$  de  $x$  dans  $K$ , un espace vectoriel réel de dimension finie  $F$  et un homéomorphisme  $\theta : \pi^{-1}(U) \rightarrow U \times F$  tel que pour tout  $x \in U$ , si  $\text{pr}_1, \text{pr}_2$  sont les projections sur les deux facteurs de  $U \times F$ , alors l’application  $\text{pr}_2 \circ \theta|_{E_x} : E_x \rightarrow F$  soit un isomorphisme linéaire et  $\text{pr}_1 \circ \theta(E_x) = \{x\}$ . Cette dernière condition signifie que le diagramme suivant est commutatif :

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\theta} & U \times F \\ \pi \searrow & & \swarrow \text{pr}_1 \\ & U & \end{array}$$

Un tel ouvert  $U$  est appelé un *ouvert trivialisant* et une telle application  $\theta$  est appelée une *trivialisant locale*. Si  $n \in \mathbb{N}$  est fixé et si les espaces vectoriels  $F$  ci-dessus sont de dimension  $n$ , nous dirons que  $E$  est de *rang*  $n$ .

Si  $\phi : K \rightarrow K'$  est une application continue entre deux espaces topologiques, un *morphisme au-dessus de  $\phi$*  d’un fibré vectoriel  $E$  sur  $K$  de projection  $\pi$ , à valeurs dans un fibré vectoriel  $E'$  sur  $K'$  de projection  $\pi'$ , est une application continue  $\Phi : E \rightarrow E'$  telle que  $\pi' \circ \Phi = \phi \circ \pi$ , ou autrement dit telle que le diagramme suivant

$$\begin{array}{ccc} E & \xrightarrow{\Phi} & E' \\ \pi \downarrow & & \downarrow \pi' \\ K & \xrightarrow{\phi} & K' \end{array}$$

soit commutatif et, telle que pour tout  $x \in K$ , l’application  $\Phi : E_x \rightarrow E'_{\phi(x)}$  soit linéaire. Un *isomorphisme* de fibrés vectoriels est un morphisme bijectif d’inverse un morphisme. Par exemple, si  $M, M'$  sont des variétés différentielles lisses, et si  $\phi : M \rightarrow M'$  est une application  $C^1$ , alors son application tangente  $T\phi : TM \rightarrow TM'$ , définie par

$$T\phi : v \mapsto \frac{d}{dt}\Big|_{t=0} \phi \circ c$$

pour n’importe quelle courbe  $c$  de classe  $C^1$  dans  $M$ , définie sur un intervalle ouvert contenant 0, de vecteur tangent  $v = \frac{d}{dt}\Big|_{t=0} c$  au temps  $t = 0$ , est un morphisme du fibré tangent  $TM$  dans le fibré tangent  $TM'$  au-dessus de  $\phi$ .

Soit  $E$  un fibré vectoriel sur un espace topologique non vide  $K$ , de projection  $\pi$ .

Un *sous-fibré* de  $E$  est une partie  $E'$  de  $E$  telle que  $\pi(E') = K$  et pour tout  $x \in K$ , l’intersection  $E_x \cap E'$  soit un sous-espace vectoriel de  $E_x$  variant continûment avec  $x$ , c’est-à-dire de sorte que  $E'$ , muni de la restriction  $\pi' = \pi|_{E'}$  et des structures de sous-espaces vectoriels sur les fibres de  $\pi'$ , soit un fibré vectoriel. Un endomorphisme de fibrés vectoriels  $\Phi : E \rightarrow E$  au-dessus d’une application continue  $\phi : K \rightarrow K$  *préserve* un sous-fibré  $E'$  de  $E$

si  $\Phi(E') \subset E'$ , ou, de manière équivalente, si pour tout  $x \in K$ , nous avons  $\Phi(E'_x) \subset E'_{\phi(x)}$ . La restriction  $\Phi|_{E'}$  est alors un endomorphisme de fibrés vectoriels au-dessus de  $\phi$ . Un fibré vectoriel  $E$  sur  $K$  est dit *somme directe* de sous-fibrés  $E_1, \dots, E_k$ , et nous noterons  $E = E_1 \oplus \dots \oplus E_k$ , si  $E_x = (E_1)_x \oplus \dots \oplus (E_k)_x$  pour tout  $x \in K$ .

Si  $K'$  est une partie de  $K$ , l'espace topologique  $E' = \pi^{-1}(K')$  muni de la restriction  $\pi' = \pi|_{E'}$  et des mêmes structures d'espaces vectoriels sur les fibres, est un fibré vectoriel sur  $K'$ , appelé la *restriction* du fibré vectoriel  $E$  à  $K'$ , et noté  $E|_{K'}$ . Par exemple, si  $M'$  est une sous-variété lisse d'une variété lisse  $M$ , alors le fibré tangent  $TM'$  de  $M'$  est un sous-fibré du fibré restriction  $TM|_{M'}$ .

Une *norme* sur  $E$  est une application continue  $\| \cdot \| : E \rightarrow \mathbb{R}$  dont la restriction  $\| \cdot \|_x$  à la fibre  $E_x$  au-dessus de tout  $x \in K$  est une norme sur l'espace vectoriel réel  $E_x$ . Un *fibré vectoriel normé* est un fibré vectoriel muni d'une norme. Par exemple, une *métrique riemannienne* sur une variété lisse  $M$  est une norme sur le fibré tangent  $TM$  de  $M$  qui est une norme euclidienne sur chaque fibre  $T_x M$  pour  $x \in M$ . Si  $E$  est muni d'une norme, nous munirons tout sous-fibré de  $E$  de la norme restreinte. Si  $K$  est métrisable, tout fibré vectoriel sur  $K$  admet une norme.<sup>211</sup> Si  $K$  est compact, par un argument de continuité et le fait que deux normes sur un espace vectoriel réel de dimension finie sont équivalentes, deux normes  $\| \cdot \|$  et  $\| \cdot \|'$  sur  $E$  sont *équivalentes*, c'est-à-dire qu'il existe  $c > 0$  tel que  $\frac{1}{c} \| \cdot \| \leq \| \cdot \|' \leq c \| \cdot \|$ .

Si  $\Phi : E \rightarrow E'$  est un morphisme de fibrés vectoriels, entre deux fibrés vectoriels normés, au-dessus d'une application continue  $\phi : K \rightarrow K'$ , sa *norme d'opérateur* est

$$\| \Phi \| = \sup_{x \in K} \| \Phi_x \|$$

où  $\| \Phi_x \| = \sup_{v \in E_x - \{0\}} \frac{\| \Phi(v) \|_{\phi(x)}}{\| v \|_x}$  est la norme d'opérateur de l'opérateur linéaire  $\Phi_x = \Phi|_{E_x}$  entre les espaces vectoriels normés  $E_x$  et  $E_{\phi(x)}$ . Le morphisme  $\Phi$  est dit *borné* si sa norme d'opérateur est finie, ce qui est par exemple le cas si  $K$  est compact.

Une *section ensembliste* (respectivement *section*) du fibré vectoriel  $E$  est une application (respectivement application continue)  $\sigma : K \rightarrow E$  telle que  $\pi \circ \sigma$  soit l'application identité sur  $K$ . Cette égalité  $\pi \circ \sigma = \text{id}$  signifie que le diagramme suivant est commutatif :

$$\begin{array}{ccc} & E & \\ \sigma \nearrow & & \downarrow \pi \\ K & \xrightarrow{\text{id}} & K \end{array}$$

Nous noterons  $\Gamma^0(E)$  l'espace vectoriel réel<sup>212</sup> des sections de  $E$ . Si  $E$  est un fibré vectoriel normé, nous noterons  $\Gamma_b^0(E)$  le sous-espace vectoriel de  $\Gamma^0(E)$  des sections bornées, muni

211. En effet, soit  $(U_i)_{i \in I}$  un recouvrement ouvert de  $K$  par des ouverts trivialisants de  $E$ , et fixons  $\theta_i : \pi^{-1}(U_i) \rightarrow U_i \times F_i$  une trivialisations locale au-dessus de  $U_i$ . Fixons une norme  $\| \cdot \|_{F_i}$  sur l'espace vectoriel réel de dimension finie  $F_i$ , et notons  $\| \cdot \|_i : v \mapsto \| \text{pr}_2 \circ \theta_i(v) \|_{F_i}$ , qui est une norme sur le fibré vectoriel  $E|_{U_i}$ . Puisque  $K$  est métrisable (hypothèse qui peut être affaiblie, mais qui sera suffisante pour nous), il existe (voir par exemple [Dug]) une partition de l'unité  $(\varphi_i)_{i \in I}$  subordonnée au recouvrement  $(U_i)_{i \in I}$ , c'est-à-dire la donnée pour tout  $i \in I$  d'une application  $\varphi_i : M \rightarrow [0, 1]$  continue et nulle en dehors de  $U_i$ , telle que la somme  $\sum_{i \in I} \varphi_i$  soit localement finie et égale à 1. Posons alors

$$\| \cdot \| = \sum_{i \in I} \varphi_i \circ \pi \| \cdot \|_i,$$

en prolongeant  $\| \cdot \|_i$  par 0 en dehors de  $\pi^{-1}(U_i)$ . Alors  $\| \cdot \|$  est une norme sur  $E$ .

212. pour les opérations points par points  $\sigma + \sigma' : x \mapsto \sigma(x) + \sigma'(x)$  et  $\lambda \sigma : x \mapsto \lambda \sigma(x)$  pour tous les

de la norme uniforme

$$\|\sigma\|_\infty = \sup_{x \in K} \|\sigma(x)\|_x .$$

Remarquons que si  $K$  est compact, alors toute section est bornée.

Notons  $\Gamma_b(E)$  l'espace vectoriel des sections ensemblistes (donc pas forcément continues) du fibré vectoriel  $E$ , qui sont bornées, muni de la norme uniforme ci-dessus.

**Proposition 10.10.** *Les espaces vectoriels normés  $\Gamma_b(E)$  et  $\Gamma_b^0(E)$  sont des espaces de Banach.*

**Démonstration.** L'argument est analogue à celui pour montrer la complétude de la norme uniforme de  $C_b^0(K; \mathbb{R})$ .

Pour tout  $x \in K$ , l'application de  $\Gamma_b(E)$  dans  $E_x$  d'évaluation en  $x$  des sections, définie par  $\sigma \mapsto \sigma(x)$ , est 1-lipschitzienne par définition de la norme uniforme. Puisque  $E_x$  est de dimension finie, sa norme  $\|\cdot\|_x$  est complète. Si  $(\sigma_n)_{n \in \mathbb{N}}$  est une suite de Cauchy dans  $\Gamma_b(E)$ , alors  $(\sigma_n(x))_{n \in \mathbb{N}}$  est une suite de Cauchy dans  $E_x$ , donc converge dans  $E_x$  vers un élément que nous notons  $\sigma(x) \in E_x$ . Soit  $\epsilon > 0$ . Soit  $N \in \mathbb{N}$  tel que si  $m, n \geq N$  alors  $\|\sigma_n - \sigma_m\|_\infty \leq \epsilon$ . Alors pour tout  $x \in K$ , nous avons  $\|\sigma_n(x) - \sigma_m(x)\|_x \leq \epsilon$ , donc en faisant tendre  $n$  vers  $+\infty$ , nous avons  $\|\sigma(x) - \sigma_m(x)\|_x \leq \epsilon$ . Par conséquent, la suite  $(\sigma_n)_{n \in \mathbb{N}}$  converge vers l'application  $\sigma : x \mapsto \sigma(x)$  pour la norme uniforme.

En particulier, la section  $\sigma$  est bornée, car toute section ensembliste à distance uniforme au plus 1 d'une section ensembliste de norme uniforme au plus  $C$  est de norme uniforme au plus  $C + 1$ .

Ceci montre que  $\Gamma_b(E)$  est un espace de Banach. Montrons que le sous-espace vectoriel  $\Gamma_b^0(E)$  est fermé dans  $\Gamma_b(E)$ . Supposons qu'une suite  $(\sigma_n)_{n \in \mathbb{N}}$  dans  $\Gamma_b^0(E)$  converge pour la norme uniforme vers  $\sigma \in \Gamma_b(E)$ . Montrons que  $\sigma$  est continue. Soit  $x_0 \in K$ . Soit  $V$  un voisinage ouvert de  $\sigma(x_0)$ , montrons qu'il existe un voisinage ouvert  $U$  de  $x_0$  tel que  $\sigma(x) \in V$  pour tout  $x \in U$ . Nous pouvons supposer que  $V \subset \pi^{-1}(U_1)$  où  $U_1$  est un voisinage ouvert de  $x_0$  trivialisant  $E$ , c'est-à-dire tel qu'il existe un espace vectoriel  $F$  et une trivialisatation locale  $\theta : \pi^{-1}(U_1) \rightarrow U_1 \times F$ . Fixons une norme  $\|\cdot\|_F$  sur  $F$ , et notons  $B(w, r)$  sa boule de centre  $w \in F$  et de rayon  $r > 0$ . Montrons que quitte à réduire  $U_1$ , il existe  $c > 0$  tel que pour tout  $v \in \pi^{-1}(U_1)$ , nous avons

$$\|\text{pr}_2 \circ \theta(v)\|_F \leq c \|v\| .$$

En effet, l'application  $f$  de  $U_1$  dans  $\mathbb{R}$  définie par  $x \mapsto \inf_{w \in F : \|w\|_F=1} \|\theta^{-1}(x, w)\|$  est continue par la continuité de  $\theta^{-1}$  et la compacité de la sphère unité de  $F$ . Elle est strictement positive en  $x_0$  puisque  $v \mapsto \|v\|_{x_0}$  est une norme sur  $E_{x_0}$ . Quitte à réduire  $U_1$ , il existe donc  $c > 0$  tel que  $f(x) \geq \frac{1}{c}$ , ce qui montre la formule centrée ci-dessus.

Par la continuité de  $\theta^{-1}$ , si  $\epsilon > 0$  est assez petit, il existe un voisinage ouvert  $U_2$  de  $x_0$  contenu dans  $U_1$  tel que  $\theta^{-1}(U_2 \times B(\text{pr}_2 \circ \theta(\sigma(x_0)), \epsilon))$  soit contenu dans  $V$ . Soit  $N \in \mathbb{N}$  tel que

$$\|\sigma_N - \sigma\|_\infty \leq \frac{\epsilon}{3c} .$$

Par la continuité de  $\sigma_N$  et de  $\theta$ , il existe un voisinage ouvert  $U_3$  de  $x_0$  contenu dans  $U_2$  tel que, pour tout  $x \in U_3$ , nous avons

$$\|\text{pr}_2 \circ \theta(\sigma_N(x)) - \text{pr}_2 \circ \theta(\sigma_N(x_0))\|_F \leq \frac{\epsilon}{3} .$$

---

$\sigma, \sigma' \in \Gamma^0(E)$  et  $\lambda \in \mathbb{R}$ , qui sont bien définies par la structure d'espace vectoriel réel de chaque fibre, et qui sont continues par la continuité de ces structures

Donc pour tout  $x \in U_3$ , nous avons  $\text{pr}_1 \circ \theta(\sigma(x)) = \pi \circ \sigma(x) = x \in U_3$  et

$$\begin{aligned} & \| \text{pr}_2 \circ \theta(\sigma(x)) - \text{pr}_2 \circ \theta(\sigma(x_0)) \|_F \leq \| \text{pr}_2 \circ \theta(\sigma(x)) - \text{pr}_2 \circ \theta(\sigma_N(x)) \|_F \\ & \quad + \| \text{pr}_2 \circ \theta(\sigma_N(x)) - \text{pr}_2 \circ \theta(\sigma_N(x_0)) \|_F + \| \text{pr}_2 \circ \theta(\sigma_N(x_0)) - \text{pr}_2 \circ \theta(\sigma(x_0)) \|_F \\ & \leq c \| \sigma(x) - \sigma_N(x) \| + \| \text{pr}_2 \circ \theta(\sigma_N(x)) - \text{pr}_2 \circ \theta(\sigma_N(x_0)) \|_F + c \| \sigma_N(x_0) - \sigma(x_0) \| \\ & \leq c \frac{\epsilon}{3c} + \frac{\epsilon}{3} + c \frac{\epsilon}{3c} = \epsilon . \end{aligned}$$

Par conséquent, pour tout  $x \in U_3$ , nous avons  $\sigma(x) \in \theta^{-1}(U_3 \times B(\text{pr}_2 \circ \theta(\sigma(x_0)), \epsilon)) \subset V$ , ce qui montre la continuité de  $\sigma$ .  $\square$

### 10.4.2 Parties hyperboliques d'une variété munie d'un difféomorphisme

Soient  $M$  une variété différentielle lisse et  $\pi : TM \rightarrow M$  son fibré tangent (de fibre au-dessus de  $x \in M$  notée  $T_x M = \pi^{-1}(x)$ ). Soient  $\phi : M \rightarrow M$  un  $C^1$ -difféomorphisme et  $T\phi : TM \rightarrow TM$  son application tangente, qui vérifie que  $T_x \phi = (T\phi)|_{T_x M}$  est un isomorphisme linéaire de  $T_x M$  dans  $T_{\phi(x)} M$ , dépendant continûment de  $x$ . Soit  $\Lambda$  une partie de  $M$  invariante par  $\phi$ , c'est-à-dire telle que  $\phi(\Lambda) = \Lambda$ . Nous munissons  $\Lambda$  de la topologie induite par  $M$ .

Nous noterons encore  $\pi : TM|_\Lambda = \bigcup_{x \in \Lambda} T_x M \rightarrow \Lambda$  la restriction de  $\pi$  à la préimage de  $\Lambda$  par  $\pi$ , qui est un fibré vectoriel sur  $\Lambda$ . Nous fixons une *norme* sur  $TM|_\Lambda$ , c'est-à-dire une application continue notée  $\| \cdot \| : TM|_\Lambda \rightarrow \mathbb{R}$  dont la restriction  $\| \cdot \|_x$  à la fibre  $T_x M$  au-dessus de tout  $x \in \Lambda$  est une norme sur l'espace vectoriel réel  $T_x M$ . Si  $\Lambda$  est compact, deux normes  $\| \cdot \|$  et  $\| \cdot \|'$  sur  $TM|_\Lambda$  sont *équivalentes*, c'est-à-dire qu'il existe  $c > 0$  tel que  $\frac{1}{c} \| \cdot \| \leq \| \cdot \|' \leq c \| \cdot \|$ . Nous munissons tout sous-fibré vectoriel de  $TM|_\Lambda$  de la norme restreinte.

Pour tous les  $x \in \Lambda$  et  $\ell \in \mathcal{L}(T_x M, T_{\phi(x)} M)$ , notons  $\| \ell \|_x = \sup_{v \in T_x M - \{0\}} \frac{\| \ell(v) \|_{\phi(x)}}{\| v \|_x}$  la norme d'opérateur de  $\ell$ . Nous supposons que

$$\sup_{x \in \Lambda} \max \{ \| T_x \phi \|_x, \| T_x(\phi^{-1}) \|_x \} < +\infty , \quad (92)$$

ce qui est automatiquement vérifié si  $\Lambda$  est compact.

Un *champ de vecteurs* sur  $\Lambda$  est une section continue de  $\pi : TM|_\Lambda \rightarrow \Lambda$ , c'est-à-dire une application continue  $X : \Lambda \rightarrow TM|_\Lambda$  telle que  $X(x) \in T_x M$  pour tout  $x \in \Lambda$ . Nous notons  $\Gamma_b^0(TM|_\Lambda)$  l'espace de Banach (voir la proposition 10.10) des champs de vecteurs  $X$  (continus) bornés (ceci est automatique par continuité si  $\Lambda$  est compact) sur  $\Lambda$  pour la norme uniforme

$$\| X \|_\infty = \sup_{x \in \Lambda} \| X(x) \|_x . \quad (93)$$

La transformation  $\phi$  induit un automorphisme linéaire continu  $\phi_*$  de l'espace de Banach  $\Gamma_b^0(TM|_\Lambda)$  en posant, pour tout champ de vecteurs  $X \in \Gamma_b^0(TM|_\Lambda)$ ,

$$\phi_*(X) : x \in \Lambda \mapsto T_{\phi^{-1}(x)} \phi (X(\phi^{-1}(x))) \in T_x M , \quad (94)$$

dont l'inverse est  $(\phi_*)^{-1} = (\phi^{-1})_*$ . Notons que l'application  $\phi_*(X) : \Lambda \rightarrow TM|_\Lambda$  est bien continue bornée par la formule (92), que  $\text{id}_* = \text{id}$ , et que si  $\psi : M \rightarrow M$  est un  $C^1$ -difféomorphisme de  $M$  tel que  $\psi(\Lambda) = \Lambda$  et  $\sup_{x \in \Lambda} \max \{ \| T_x \psi \|_x, \| T_x(\psi^{-1}) \|_x \} < +\infty$ , alors

$$(\psi \circ \phi)_* = \psi_* \circ \phi_* .$$



En particulier, pour tout  $n \in \mathbb{Z}$ , nous avons  $(\phi^n)_* = (\phi_*)^n$ .

Notons  $\Gamma_b(TM|_\Lambda)$  l'espace de Banach (voir la proposition 10.10) des champs de vecteurs pas forcément continus sur  $\Lambda$  (c'est-à-dire des applications pas forcément continues notées  $\sigma : \Lambda \rightarrow TM|_\Lambda$  telles que  $\sigma(x) \in T_x M$  pour tout  $x \in \Lambda$ ), mais qui sont bornés, muni de la norme uniforme  $\|\sigma\|_\infty = \sup_{x \in \Lambda} \|\sigma(x)\|_x$  définie par la même formule (93) que ci-dessus, et de l'automorphisme linéaire continu

$$\phi_* : \sigma \mapsto \left\{ x \mapsto T_{\phi^{-1}(x)} \phi \left( \sigma(\phi^{-1}(x)) \right) \right\} \quad (95)$$

défini par la même formule (94) que ci-dessus. En particulier,  $\Gamma_b^0(TM|_\Lambda)$  est un sous-espace vectoriel fermé de  $\Gamma_b(TM|_\Lambda)$ .

**Définition 10.11.** *Soit  $\Lambda$  une partie de  $M$  invariante par  $\phi$ , munie d'une norme sur  $TM|_\Lambda$ . Nous dirons que  $\Lambda$  est une partie hyperbolique de  $M$  pour  $\phi$  si la propriété (92) est vérifiée et s'il existe  $\kappa_s, \kappa_u \in \mathbb{R}$  tels que  $0 < \kappa_s < 1 < \kappa_u$  (nous dirons alors que  $\Lambda$  est une partie  $(\kappa_s, \kappa_u)$ -hyperbolique de  $M$  lorsqu'il convient de préciser), de sorte que l'une des trois conditions équivalentes suivantes soit vérifiée :*

- (1) *l'automorphisme linéaire continu  $\phi_*$  de l'espace de Banach  $\Gamma_b^0(TM|_\Lambda)$  est  $(\kappa_s, \kappa_u)$ -hyperbolique,*
- (2) *l'automorphisme linéaire continu  $\phi_*$  de l'espace de Banach  $\Gamma_b(TM|_\Lambda)$  est  $(\kappa_s, \kappa_u)$ -hyperbolique,*
- (3) *il existe une décomposition  $TM|_\Lambda = E^s \oplus E^u$  en somme directe de deux sous-fibrés  $E^s$  et  $E^u$  sur  $\Lambda$  continus et  $T\phi$ -invariants (c'est-à-dire  $T\phi(E^s) = E^s$  et  $T\phi(E^u) = E^u$ )<sup>213</sup>, appelés les sous-fibrés stable et instable de  $\phi$ , et des constantes  $c > 0$  et  $\kappa'_s, \kappa'_u \in \mathbb{R}$  telles que  $0 < \kappa'_s < \kappa_s < 1 < \kappa_u < \kappa'_u$ , de sorte que, pour tout  $n \in \mathbb{N}$*

$$\|T\phi^n|_{E^s}\| \leq c \kappa'_s{}^n, \quad \|T\phi^{-n}|_{E^u}\| \leq c \kappa'_u{}^{-n}.$$

**Démonstration de l'équivalence des définitions.** Nous renvoyons à la partie 10.11 en appendice à ce chapitre 10 pour des rappels de théorie spectrale.

Comme  $\Gamma_b^0(TM|_\Lambda)$  est un sous-espace vectoriel fermé de  $\Gamma_b(TM|_\Lambda)$  invariant par  $\phi_*$ , le spectre de  $\phi_*$  dans  $\Gamma_b^0(TM|_\Lambda)$  est contenu dans le spectre de  $\phi_*$  dans  $\Gamma_b(TM|_\Lambda)$ . Donc si le spectre de  $\phi_*$  dans  $\Gamma_b(TM|_\Lambda)$  ne rencontre pas l'anneau  $\{z \in \mathbb{C} : \kappa_s \leq |z| \leq \kappa_u\}$ , alors le spectre de  $\phi_*$  dans  $\Gamma_b^0(TM|_\Lambda)$  non plus. Le fait que l'assertion (2) implique l'assertion (1) découle alors de l'équivalence dans la seconde partie de la proposition 10.2.

Si l'assertion (3) est vérifiée, alors la décomposition en somme directe

$$\Gamma_b(TM|_\Lambda) = \Gamma_b(E^s) \oplus \Gamma_b(E^u),$$

définie par  $\sigma \mapsto \sigma_s + \sigma_u$  où  $\sigma_s(x) \in E_x^s$ ,  $\sigma_u(x) \in E_x^u$  et  $\sigma(x) = \sigma_s(x) + \sigma_u(x)$  pour tout  $x \in \Lambda$ , vérifie les axiomes (1) à (3) du début la partie 10.1 définissant le fait que l'automorphisme linéaire continu  $\phi_*$  de  $\Gamma_b(TM|_\Lambda)$  (donné par la formule (95)) soit  $(\kappa_s, \kappa_u)$ -hyperbolique<sup>214</sup>. Donc l'assertion (2) est vérifiée.

213. Voir le lemme 10.13 ci-dessous pour un affaiblissement de ces deux conditions

214. en utilisant pour la vérification du troisième axiome susdit le fait que pour tout  $n \in \mathbb{Z}$ , nous ayons  $(\phi^n)_* = (\phi_*)^n$

Pour montrer que l'assertion (1) implique l'assertion (3), nous allons utiliser le *lemme de tensorialité* suivant, pour lequel nous renvoyons par exemple à [Pau5, Lem. 2.24]. Notons que l'espace de Banach

$$\mathcal{E} = \Gamma_b^0(TM|_\Lambda)$$

admet une structure naturelle de  $C_b^0(\Lambda; \mathbb{R})$ -module, en posant, pour tout champ de vecteurs continu borné  $X \in \mathcal{E}$  et toute application continue bornée  $f \in C_b^0(\Lambda; \mathbb{R})$ ,

$$fX : x \mapsto f(x)X(x).$$

**Lemme 10.12.** <sup>215</sup> *Soit  $A : \mathcal{E} \rightarrow \mathcal{E}$  une application continue  $C_b^0(\Lambda; \mathbb{R})$ -linéaire. Alors il existe une unique famille continue  $(A_x)_{x \in \Lambda} \in (\mathcal{L}(T_x M))_{x \in \Lambda}$  telle que, pour tous les  $x \in \Lambda$  et  $X \in \mathcal{E}$ , nous ayons*

$$(AX)(x) = A_x(X(x)). \quad \square$$

Supposons que  $\phi_* : \mathcal{E} \rightarrow \mathcal{E}$  soit  $(\kappa_s, \kappa_u)$ -hyperbolique, et soit  $\mathcal{E} = \mathcal{E}^s \oplus \mathcal{E}^u$  la décomposition en somme directe de sous-espaces de Banach donnée par la définition au début de la section 10.1 et le lemme 10.1 (2). Notons  $\pi^s : \mathcal{E} \rightarrow \mathcal{E}^s$  et  $\pi^u : \mathcal{E} \rightarrow \mathcal{E}^u$  les projections linéaires associées à cette décomposition. Remarquons que pour tous les  $X \in \mathcal{E}$  et  $f \in C_b^0(\Lambda; \mathbb{R})$ , par la formule (94) et la linéarité des applications tangentes, nous avons

$$\phi_*(fX) = f \circ \phi^{-1} \phi_*(X).$$

Donc  $(\phi_*)^n(f\pi^s(X)) = f \circ \phi^{-n} (\phi_*)^n(\pi^s(X))$  et  $(\phi_*)^{-n}(f\pi^u(X)) = f \circ \phi^n (\phi_*)^{-n}(\pi^u(X))$  pour tout  $n \in \mathbb{N}$ . Puisque les applications  $f \circ \phi^n$  pour  $n \in \mathbb{Z}$  sont uniformément bornées, les propriétés de contraction/dilatation de  $\phi_*$  (voir aussi les formules (90) et (91)) montrent que nous avons  $f\pi^s(X) \in \mathcal{E}^s$  et  $f\pi^u(X) \in \mathcal{E}^u$ . Puisque  $fX = f\pi^s(X) + f\pi^u(X)$ , ceci montre que les applications continues  $\pi^s$  et  $\pi^u$  sont  $C_b^0(\Lambda; \mathbb{R})$ -linéaires. Par le lemme ci-dessus, il existe donc des uniques familles continues  $(p_x^s)_{x \in \Lambda}$  et  $(p_x^u)_{x \in \Lambda}$  avec  $p_x^s, p_x^u \in \mathcal{L}(T_x M)$  pour tout  $x \in \Lambda$ , telles que

$$\forall X \in \mathcal{E}, \forall x \in \Lambda, \quad (\pi^s X)(x) = p_x^s(X(x)) \quad \text{et} \quad (\pi^u X)(x) = p_x^u(X(x)).$$

215. Par souci de complétude, voici une démonstration. Montrons d'abord l'unicité. En effet, pour tous les  $x \in \Lambda$  et  $v \in T_x M$ , il existe un champ de vecteurs  $X_v \in \mathcal{E}$  tel que  $X_v(x) = v$  (en construisant  $X_v$  en tout point d'une carte locale de  $M$  en  $x$ , puis en multipliant par une fonction plateau valant 1 en  $x$  pour étendre à tout  $M$ , puis en restreignant à  $\Lambda$ ). Donc si une famille  $(A_x)_{x \in \Lambda}$  vérifie l'énoncé du lemme, nous avons  $A_x(v) = (AX_v)(x)$  pour tous les  $x \in \Lambda$  et  $v \in T_x M$ , ce qui caractérise la famille  $(A_x)_{x \in \Lambda}$ .

Montrons que pour tous les  $X \in \mathcal{E}$  et  $x \in \Lambda$ , la valeur de  $AX$  en  $x$  ne dépend que de la valeur  $v = X(x)$  de  $X$  en  $x$ . En la notant  $A_x(v)$ , il est alors élémentaire de vérifier que la famille  $(A_x)_{x \in \Lambda}$  convient.

- Montrons que si  $X$  s'annule sur l'intersection avec  $\Lambda$  d'un voisinage ouvert  $U$  de  $x$  dans  $M$ , alors  $AX$  s'annule sur  $\Lambda \cap U$ . En effet, pour tout  $x \in \Lambda \cap U$ , soit  $\varphi \in C^\infty(M)$  à support compact contenu dans  $U$ , valant 1 en  $x$ . Alors  $f = \varphi|_\Lambda$  appartient à  $C_b^0(\Lambda; \mathbb{R})$  et  $fX$  est le champ de vecteurs nul sur  $\Lambda$ . Donc par la  $C_b^0(\Lambda; \mathbb{R})$ -linéarité de  $A$ , nous avons  $AX(x) = f(x)AX(x) = A(fX)(x) = 0$ .

- Montrons que si  $X$  s'annule en  $x$ , alors  $AX$  s'annule en  $x$ . En effet, considérons une suite  $(\varphi_n)_{n \in \mathbb{N}}$  dans  $C^\infty(M)$ , dont chaque élément vaut 1 sur un voisinage ouvert de  $x$  et qui est à valeurs dans  $[0, 1]$ , telle que le support  $K_n$  de  $\varphi_n$  soit compact, vérifiant  $K_{n+1} \subset K_n$  et  $\bigcap_n K_n = \{x\}$ . Alors  $f_n = \varphi_n|_\Lambda$  appartient à  $C_b^0(\Lambda; \mathbb{R})$ , le champ de vecteurs  $X - f_n X$  s'annule sur un voisinage de  $x$  et  $(f_n X)_{n \in \mathbb{N}}$  converge vers 0 dans l'espace de Banach  $\mathcal{E}$  (car  $X(x) = 0$ ). Donc, par le point précédent et par la linéarité et la continuité de  $A$ , nous avons

$$AX(x) = \lim_{n \rightarrow +\infty} (A(X - f_n X)(x) + A(f_n X)(x)) = \lim_{n \rightarrow +\infty} A(f_n X)(x) = 0.$$

Maintenant, pour tout  $x \in \Lambda$ , si  $X, X' \in \mathcal{E}$  vérifient  $X(x) = X'(x)$ , alors  $(X - X')(x) = 0$ , donc nous avons  $A(X - X')(x) = 0$ , c'est-à-dire  $AX(x) = AX'(x)$  par la linéarité de  $A$ . C'est ce que nous voulions montrer.

Par unicité, puisque  $\pi^s + \pi^u = \text{id}$ ,  $(\pi^s)^2 = \pi^s$  et  $(\pi^u)^2 = \pi^u$ , pour tout  $x \in \Lambda$ , les deux opérateurs  $p_x^s, p_x^u \in \mathcal{L}(T_x M)$  sont des projecteurs linéaires  $((p_x^s)^2 = p_x^s$  et  $(p_x^u)^2 = p_x^u)$  tels que  $p_x^s + p_x^u = \text{id}$ .

Posons  $E_x^s = p_x^s(T_x M)$  et  $E_x^u = p_x^u(T_x M)$ . Alors  $T_x M = E_x^s \oplus E_x^u$ , et par la continuité de  $(p_x^s)_{x \in \Lambda}$  et  $(p_x^u)_{x \in \Lambda}$ , les ensembles  $E^s = \bigcup_{x \in \Lambda} E_x^s$  et  $E^u = \bigcup_{x \in \Lambda} E_x^u$  sont des sous-fibrés vectoriels de  $TM|_\Lambda$  en somme directe, qui vérifient l'assertion (3).  $\square$

**Remarques.** (1) Si  $\Lambda$  est compact, la définition 10.11 est indépendante du choix de la norme sur  $TM|_\Lambda$ .

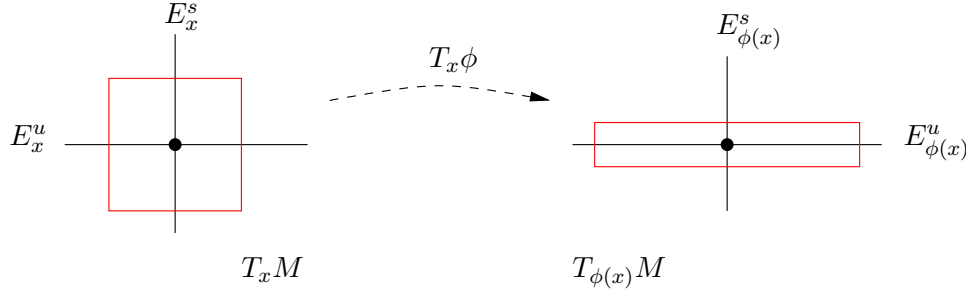
(2) Comme dans le cas linéaire, une décomposition  $TM|_\Lambda = E^s \oplus E^u$  en somme directe de deux sous-fibrés vectoriels de  $TM|_\Lambda$ , comme dans la définition 10.11, est alors unique. <sup>216</sup>

(3) Une norme  $\|\cdot\|'$  sur  $TM|_\Lambda$  est dite *adaptée* à  $\phi$  si elle est équivalente à la norme  $\|\cdot\|$  de  $TM|_\Lambda$  (ce qui est automatique si  $\Lambda$  est compact) et si

$$\forall x \in \Lambda, \forall v_s \in E_x^s, \forall v_u \in E_x^u, \quad \|v_s + v_u\|' = \max\{\|v_s\|', \|v_u\|'\}$$

et si les normes d'opérateur des restrictions à  $E_x^s$  de  $T_x \phi$  et à  $E_x^u$  de  $T_x(\phi^{-1})$  vérifient

$$\|T\phi|_{E^s}\|'_\infty = \sup_{x \in \Lambda} \|T_x \phi|_{E_x^s}\|' \leq \kappa_s, \quad \|(T\phi|_{E^u})^{-1}\|'_\infty = \sup_{x \in \Lambda} \|(T_x \phi|_{E_x^u})^{-1}\|' \leq \kappa_u^{-1}.$$



Par une démonstration analogue à celle du cas linéaire (voir le lemme 10.1 (3)), il existe au moins une norme adaptée. <sup>217</sup>

(4) Il est possible dans la définition 10.11 (voir par exemple [Yoc2, §1.2]) de remplacer  $(\Lambda, TM|_\Lambda)$  par  $(\Lambda, E)$  où  $\Lambda$  est un espace topologique, et  $\pi : E \rightarrow \Lambda$  un fibré vectoriel

216. En effet, la même démonstration que celle du lemme 10.1 (1) montre qu'en posant  $\lambda = \max\{\kappa_s, \frac{1}{\kappa_u}\}$  et avec  $c > 0$  comme dans la condition (3) de la définition 10.11, pour tout  $x \in \Lambda$ , nous avons

$$E_x^s = \{v \in T_x M : \forall n \in \mathbb{N}, \quad \|T_x \phi^n(v)\| \leq c \lambda^n \|v\|\},$$

et

$$E_x^u = \{v \in T_x M : \forall n \in \mathbb{N}, \quad \|v\| \leq c \lambda^n \|T_x \phi^n(v)\|\}.$$

217. Il suffit en effet de poser, pour  $N$  assez grand, pour tous les  $x \in \Lambda$ ,  $v_s \in E_x^s$  et  $v_u \in E_x^u$ ,

$$\|v_s + v_u\|' = \max \left\{ \sum_{i=0}^{N-1} \kappa_s^{-i} \|T_x \phi^i(v_s)\|, \sum_{i=0}^{N-1} \kappa_u^i \|T_x(\phi^{-i})(v_u)\| \right\}.$$

L'hypothèse (92) que  $\sup_{x \in \Lambda} \max\{\|T_x \phi\|_x, \|T_x(\phi^{-1})\|_x\} < +\infty$  montre que cette norme  $\|\cdot\|'$  est équivalente à  $\|\cdot\|$ .

normé au-dessus de  $\Lambda$ , et  $(\phi, T\phi)$  par  $(\phi, \Phi)$  où  $\phi : \Lambda \rightarrow \Lambda$  est un homéomorphisme et  $\Phi : E \rightarrow E$  un isomorphisme de fibrés vectoriels au-dessus de  $\phi$  tel que

$$\sup_{x \in \Lambda} \max \{ \|\Phi|_{\pi^{-1}(x)}\|, \|\Phi^{-1}|_{\pi^{-1}(x)}\| \} < +\infty .$$

Voir la sous-partie 10.4.1 pour des définitions.

(5) En fait, la condition de continuité des sous-fibrés stable et instable dans la définition 10.11 est automatique, par le lemme suivant.

**Lemme 10.13.** *Soit  $\Lambda$  une partie de  $M$  invariante par  $\phi$ , munie d'une norme sur le fibré vectoriel  $TM|_{\Lambda}$  vérifiant la propriété (92). Supposons qu'il existe  $c > 0$  et  $\lambda \in ]0, 1[$  tels que pour tout  $x \in \Lambda$ , nous ayons une décomposition en somme directe vectorielle  $T_x M = E_x^s \oplus E_x^u$ , préservée par  $T_x \phi$  (c'est-à-dire  $T_x \phi(E_x^s) \subset E_{\phi(x)}^s$  et  $T_x \phi(E_x^u) \subset E_{\phi(x)}^u$ ) telle que pour tous les  $n \in \mathbb{N}$ ,  $v \in E_x^s$  et  $v' \in E_x^u$ , nous ayons*

$$\|T_x \phi^n(v)\| \leq c \lambda^n \|v\|, \quad \|v'\| \leq c \lambda^n \|T_x \phi^n v'\|. \quad (96)$$

Alors les dimensions de  $E_x^s$  et  $E_x^u$  sont localement constantes, les sous-espaces vectoriels  $E_x^s$  et  $E_x^u$  varient continûment en  $x$ , et  $\Lambda$  est une partie hyperbolique de  $M$  pour  $\phi$ .

**Démonstration.** Puisque  $T_x \phi$  est un isomorphisme linéaire de  $T_x M$  dans  $T_{\phi(x)} M$ , les conditions  $T_x \phi(E_x^s) \subset E_{\phi(x)}^s$  et  $T_x \phi(E_x^u) \subset E_{\phi(x)}^u$  (et un argument de blocs) impliquent que  $T_x \phi(E_x^s) = E_{\phi(x)}^s$  et  $T_x \phi(E_x^u) = E_{\phi(x)}^u$ , et en particulier  $T_x \phi|_{E_x^u}$  est inversible.

Soient  $(x_k)_{k \in \mathbb{N}}$  et  $(w_k)_{k \in \mathbb{N}}$  des suites dans  $\Lambda$  et  $TM$ , convergeant vers  $x \in \Lambda$  et  $v \in TM$  respectivement. Montrons, par une démonstration analogue<sup>218</sup> à celle du lemme 10.1, que si  $w_k \in E_{x_k}^s$  pour tout  $k \in \mathbb{N}$ , alors nous avons  $v \in E_x^s$ .

Le vecteur  $v \in T_x M$  s'écrit de manière unique  $v = v_s + v_u$ , avec  $v_s \in E_x^s$  et  $v_u \in E_x^u$ . Pour tout  $n \in \mathbb{N}$ , nous avons  $T_x \phi^n(v_u) \in E_{\phi^n(x)}^u$ . Donc par la continuité de l'application tangente et de la norme, puisque  $w_k \in E_{x_k}^s$  et  $v_s \in E_x^s$ , pour tout  $n \in \mathbb{N}$ , nous avons

$$\begin{aligned} \|v_u\| &= \|T_{\phi^n(x)} \phi^{-n}(T_x \phi^n(v_u))\| \leq c \lambda^n \|T_x \phi^n(v_u)\| = c \lambda^n \|T_x \phi^n(v - v_s)\| \\ &\leq c \lambda^n (\|T_x \phi^n(v)\| + \|T_x \phi^n(v_s)\|) = c \lambda^n \left( \lim_{k \rightarrow +\infty} \|T_{x_k} \phi^n(w_k)\| + \|T_x \phi^n(v_s)\| \right) \\ &\leq c^2 \lambda^{2n} \left( \lim_{k \rightarrow +\infty} \|w_k\| + \|v_s\| \right) = c^2 \lambda^{2n} (\|v\| + \|v_s\|) . \end{aligned}$$

Comme  $c^2 \lambda^{2n}$  tend vers 0 quand  $n$  tend vers  $+\infty$ , ceci implique que  $v_u = 0$  donc que  $v = v_s$  appartient à  $E_x^s$ .

Une démonstration analogue montre si  $w_k \in E_{x_k}^u$  pour tout  $k \in \mathbb{N}$ , alors nous avons  $v \in E_x^u$ .

En travaillant dans un voisinage compact de  $x$  (contenant les  $x_k$ , quitte à extraire) contenu dans un domaine de carte locale, nous pouvons supposer que la dimension de  $M$  est constante, égale à  $N$ , et que la norme de  $TM|_{\Lambda}$  est euclidienne sur chaque fibre. Quitte à extraire, nous pouvons supposer que la suite  $(\dim E_{x_k}^s)_{k \in \mathbb{N}}$  dans  $\{0, \dots, N\}$  est constante, de valeur notée  $r$ . Pour tout  $k \in \mathbb{N}$ , soit  $\mathcal{B}_k$  une base de  $T_{x_k} M$ , concaténation d'une base

218. et en fait conséquence de cette démonstration, car la dernière assertion du lemme 10.13 découle aussi du fait que les hypothèses du lemme 10.13 font que la décomposition en somme directe  $\Gamma_b(TM|_{\Lambda}) = \Gamma_b(E^s) \oplus \Gamma_b(E^u)$  vérifie la définition 10.11 (2).

orthonormée  $\mathcal{B}_k^s$  de  $E_{x_k}^s$  et d'une base orthonormée  $\mathcal{B}_k^u$  de  $E_{x_k}^u$ . Quitte à extraire, la suite  $(\mathcal{B}_k)_{k \in \mathbb{N}}$  converge vers une suite finie de  $T_x M$ , formée, par l'étude de limite précédente, d'une partie libre  $\mathcal{B}^s$  de  $r$  éléments de  $E_x^s$  et d'une partie libre  $\mathcal{B}^u$  de  $N - r$  éléments de  $E_x^u$ . Puisque  $T_x M = E_x^s \oplus E_x^u$ , ceci implique que  $\mathcal{B}^s$  est une base de  $E_x^s$  et  $\mathcal{B}^u$  une base de  $E_x^u$ . Le résultat en découle.  $\square$

**Exemples.** (1) Définissons une orbite périodique *hyperbolique* comme l'orbite d'un point périodique  $x$  de période  $p \geq 1$  tel que l'application linéaire  $T_x(\phi^p)$ , qui va de l'espace vectoriel de dimension finie  $T_x M$  dans  $T_{\phi^p(x)} M = T_x M$ , soit un automorphisme linéaire hyperbolique de  $T_x M$ . Il est immédiat qu'une orbite périodique hyperbolique est une partie hyperbolique de  $M$  pour  $\phi$ , en posant, avec  $T_x M = E_x^s \oplus E_x^u$  la décomposition en sous-espace stable et instable de  $T_x M$  pour l'endomorphisme linéaire hyperbolique  $T_x(\phi^p)$  et pour tout  $k = 1, \dots, p - 1$ ,

$$E_{\phi^k(x)}^s = T_x(\phi^k)(E_x^s) \quad \text{et} \quad E_{\phi^k(x)}^u = T_x(\phi^k)(E_x^u).$$

(2) Soit  $M$  une variété différentielle lisse, dont le fibré tangent est muni d'une norme (par exemple riemannienne, indifférente lorsque  $M$  est compacte, voir les rappels de la partie 10.4.1). Un  $C^1$ -difféomorphisme  $\phi : M \rightarrow M$  est appelé un *difféomorphisme Anosov* de  $M$  si la variété  $M$  toute entière est une partie hyperbolique pour  $\phi$ .

Par exemple, si la variété est le tore  $\mathbb{T}^N$  pour un  $N \in \mathbb{N} - \{0\}$ , et si  $\mathcal{M} \in \text{GL}_N(\mathbb{Z})$  est une matrice entière inversible, n'ayant pas de valeur propre de module 1, alors la transformation  $\phi_{\mathcal{M}} : \mathbb{T}^N \rightarrow \mathbb{T}^N$  définie dans le chapitre 4 est un difféomorphisme Anosov. En effet,  $\mathcal{M}$  (identifié à la transformation linéaire de  $\mathbb{R}^N$  de matrice  $\mathcal{M}$ ) est un automorphisme linéaire hyperbolique de  $\mathbb{R}^N$  (voir la proposition 10.2). Notons  $\mathbb{R}^N = E^s \oplus E^u$  la décomposition de  $\mathbb{R}^N$  en somme directe invariante par  $\mathcal{M}$  telle que les valeurs propres complexes de  $\mathcal{M}|_{E^s}$  et  $\mathcal{M}|_{E^u}$  soient respectivement de module strictement inférieur à 1 et strictement supérieur à 1. Alors le fibré tangent  $T\mathbb{T}^N$  s'identifie à  $\mathbb{T}^N \times \mathbb{R}^N$ , la projection  $\pi : T\mathbb{T}^N \rightarrow \mathbb{T}^N$  s'identifie à la première projection  $\text{pr}_1 : (\mathbb{T}^N \times \mathbb{R}^N) \rightarrow \mathbb{T}^N$ , la différentielle  $T\phi_{\mathcal{M}}$  s'identifie à l'application  $\phi_{\mathcal{M}} \times \mathcal{M} : (x, v) \mapsto (\phi_{\mathcal{M}}(x), \mathcal{M}v)$ , et les sous-fibrés stable et instable de  $T\mathbb{T}^N$  pour  $\phi_{\mathcal{M}}$  s'identifient respectivement à  $\mathbb{T}^N \times E^s$  et  $\mathbb{T}^N \times E^u$ . Le caractère Anosov du difféomorphisme  $\phi_{\mathcal{M}}$  est au coeur de la démonstration de sa stabilité structurelle (voir la partie 10.3).

(3) Soit  $M$  une variété différentielle lisse, dont le fibré tangent est muni d'une norme (par exemple riemannienne, indifférente lorsque  $M$  est compacte, voir la partie 10.4.1). Comme introduit par Smale [Sma], un  $C^1$ -difféomorphisme  $\phi : M \rightarrow M$  est dit *Axiome A* si

- l'ensemble non errant  $\Omega(\phi)$  de  $\phi$  est une partie hyperbolique de  $M$  pour  $\phi$ ,
- l'ensemble non errant  $\Omega(\phi)$  est l'adhérence  $\overline{\text{Per}(\phi)}$  de l'ensemble des points périodiques de  $\phi$ .

**Exercice E.57.** *Montrer que si  $\mathcal{M} \in \text{GL}_N(\mathbb{Z})$  est une matrice entière inversible, sans valeur propre de module 1, alors la transformation  $\phi_{\mathcal{M}}$  est Axiome A.*

(4) Il existe une définition analogue à celle de l'exemple (2) pour les systèmes dynamiques à temps continu. Soit  $M$  une variété différentielle lisse, dont le fibré tangent est muni

d'une norme (par exemple riemannienne, indifférente lorsque  $M$  est compacte, voir la partie 10.4.1). Soit  $(\phi^t)_{t \in \mathbb{R}}$  un groupe à un paramètre de classe  $C^1$  de  $C^1$ -difféomorphismes <sup>219</sup> de  $M$  tel que

$$\sup_{x \in M, t \in [-1, 1]} \|T_x \phi^t\|_x < +\infty,$$

ce qui automatiquement vérifié si la variété  $M$  est compacte. Alors  $(\phi^t)_{t \in \mathbb{R}}$  est appelé un *flot d'Anosov* s'il existe une décomposition

$$TM = E^0 \oplus E^s \oplus E^u$$

en somme directe de trois sous-fibrés vectoriels  $E^0$ ,  $E^s$  et  $E^u$  sur  $M$  continus et  $T\phi$ -invariants (c'est-à-dire tels que  $T\phi^t(E^0) = E^0$ ,  $T\phi^t(E^s) = E^s$  et  $T\phi^t(E^u) = E^u$  pour tout  $t \in \mathbb{R}$ ), appelés les *sous-fibrés neutre, stable* et *instable* de  $\phi$ , et des constantes  $c > 0$  et  $\lambda \in ]0, 1[$  telles que, pour tout  $x \in M$ ,

$$E_x^0 = \mathbb{R} \frac{d}{dt} \Big|_{t=0} \phi^t x$$

et, pour tout  $t \geq 0$ , les normes d'opérateur des restrictions de  $T\phi^t$  à  $E^s$  et de  $T\phi^{-t}$  à  $E^u$  vérifient

$$\|T\phi^t|_{E^s}\| \leq c\lambda^t, \quad \|T\phi^{-t}|_{E^u}\| \leq c\lambda^t.$$

Comme pour les difféomorphismes d'Anosov, il n'est pas nécessaire de demander la continuité des sous-fibrés  $E^0$ ,  $E^s$ ,  $E^u$ , elle découle des autres propriétés, et il suffit de vérifier que  $T\phi^t(E^0) \subset E^0$ ,  $T\phi^t(E^s) \subset E^s$  et  $T\phi^t(E^u) \subset E^u$ .

Par exemple, si  $\Gamma$  est un sous-groupe discret de  $\mathrm{PSL}_2(\mathbb{R})$ , alors le flot géodésique  $(\mathfrak{g}^t)_{t \in \mathbb{R}}$  sur la variété quotient <sup>220</sup>  $M = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  de dimension 3 est un flot d'Anosov, dont le sous-fibré stable (respectivement instable) est le fibré tangent aux orbites du flot horocyclique stable (respectivement instable). En effet, pour tout  $v \in M$  (c'est l'orbite par  $\Gamma$  d'un vecteur tangent unitaire à  $\mathbb{H}_{\mathbb{R}}^2$ ), posons

$$X_v^0 = \frac{d}{d\tau} \Big|_{\tau=0} \mathfrak{g}^\tau v, \quad X_v^s = \frac{d}{d\tau} \Big|_{\tau=0} \mathfrak{h}^\tau v, \quad X_v^u = \frac{d}{d\tau} \Big|_{\tau=0} \bar{\mathfrak{h}}^\tau v.$$

Nous munissons le fibré vectoriel  $TM$  de la norme (dite *norme de Sasaki*), qui est euclidienne sur chaque fibre et rend la base  $(X_v^0, X_v^s, X_v^u)$  de  $T_v M$  orthonormée pour tout  $v \in M$ .

Alors par les relations de commutations

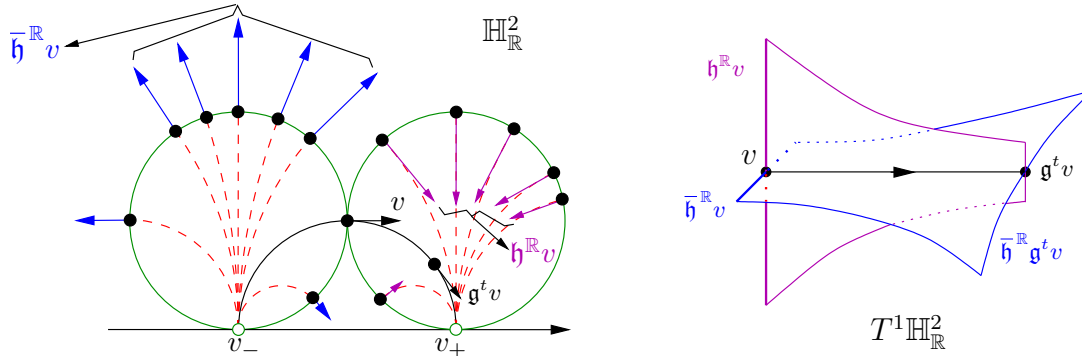
$$\forall t, \tau \in \mathbb{R}, \quad \mathfrak{g}^t \circ \mathfrak{g}^\tau = \mathfrak{g}^\tau \circ \mathfrak{g}^t, \quad \mathfrak{g}^t \circ \mathfrak{h}^\tau = \mathfrak{h}^\tau e^{-t} \circ \mathfrak{g}^t \quad \text{et} \quad \mathfrak{g}^t \circ \bar{\mathfrak{h}}^\tau = \bar{\mathfrak{h}}^\tau e^t \circ \mathfrak{g}^t$$

(voir l'exercice E.54 (1) et (2)) et par le théorème de dérivation des fonctions composées, nous avons

$$\forall t \in \mathbb{R}, \quad T_x \mathfrak{g}^t(X_v^0) = X_{\mathfrak{g}^t v}^0, \quad T_x \mathfrak{g}^t(X_v^s) = e^{-t} X_{\mathfrak{g}^t v}^s, \quad T_x \mathfrak{g}^t(X_v^u) = e^t X_{\mathfrak{g}^t v}^u.$$

219. Ceci signifie que l'application  $\phi : \mathbb{R} \times M \rightarrow M$  définie par  $(t, x) \mapsto \phi^t(x)$  est de classe  $C^1$  (outre que  $\phi^0 = \mathrm{id}$  et  $\phi^{t+\tau} = \phi^t \circ \phi^\tau$  pour tous les  $t, \tau \in \mathbb{R}$ ).

220. Comme vu dans la partie 9.5, le fibré tangent unitaire  $T^1\mathbb{H}_{\mathbb{R}}^2$  s'identifie avec  $G = \mathrm{PSL}_2(\mathbb{R})$ . Puisque le sous-groupe  $\Gamma$  est discret dans  $G$ , l'action par translations à gauche de  $\Gamma$  sur  $G$  est propre et libre. Donc la projection canonique  $T^1\mathbb{H}_{\mathbb{R}}^2 = G \rightarrow M = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2 = \Gamma \backslash G$  est un revêtement. La structure de *variété quotient* sur  $M$  est alors l'unique structure de variété lisse telle que cette projection canonique soit une submersion lisse. Voir par exemple [Pau2, §2.4.2] pour des explications.



Ceci montre que  $(\mathbf{g}^t)_{t \in \mathbb{R}}$  est un flot d'Anosov, avec décomposition  $TM = E^0 \oplus E^s \oplus E^u$  en somme directe de sous-fibrés neutre, stable et instable (tous trois de fibres de dimension 1) donnée par

$$\forall v \in M, \quad E_v^0 = \mathbb{R} X_v^0, \quad E_v^s = \mathbb{R} X_v^s, \quad E_v^u = \mathbb{R} X_v^u,$$

et avec pour norme adaptée à  $(\mathbf{g}^t)_{t \in \mathbb{R}}$  sur  $TM$

$$\forall v \in M, \quad \left\| x_0 X_v^0 + x_s X_v^s + x_u X_v^u \right\|'_v = \max\{|x_0|, |x_s|, |x_u|\}.$$

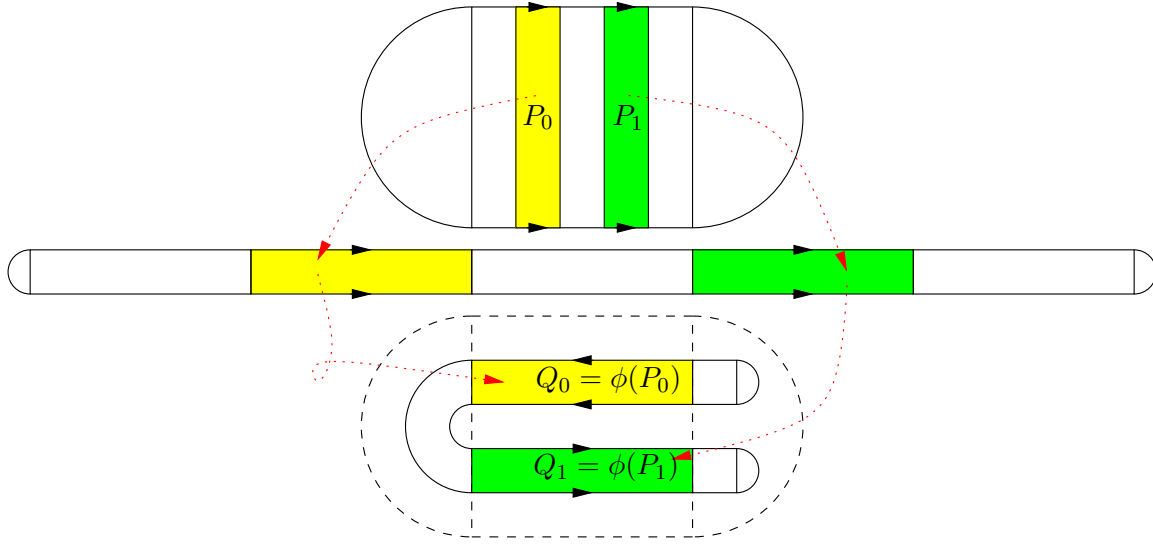
Remarquons que si  $\Gamma$  est trivial, alors l'ensemble non errant de  $(\mathbf{g}^t)_{t \in \mathbb{R}}$  est vide. Donc la dynamique du flot géodésique  $(\mathbf{g}^t)_{t \in \mathbb{R}}$ , même s'il est d'Anosov, ne sera intéressante que lorsque le groupe  $\Gamma$  est intéressant (mais pas forcément de quotient  $M = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$  compact, comme par exemple pour le groupe modulaire  $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ , voir la partie 9.5). Nous renvoyons par exemple au joli livre [Dal] pour les aspects de dynamique topologique du flot géodésique sur les surfaces hyperboliques.

## 10.5 Le fer à cheval de Smale

Je recommande de visualiser le chapitre 6 du superbe film d'animation **Chaos** par Leys-Ghys-Alvarez sur

<https://www.chaos-math.org/fr.html>

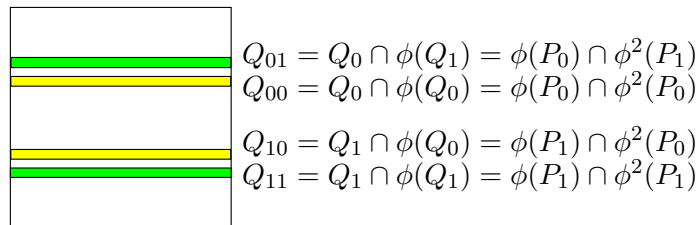
Le *fer à cheval*, introduit par Smale (voir par exemple [Sma]), est l'un des exemples les plus simples de difféomorphisme admettant un ensemble hyperbolique remarquable (en particulier d'intérieur vide, contrairement au cas des difféomorphismes d'Anosov), qui sert d'illustration à de très nombreux phénomènes de dynamique hyperbolique (dont la propriété d'être Axiome A, d'avoir un codage par un décalage de Bernoulli, d'être un modèle pour des points homoclines, etc). Cet exemple est traité dans de nombreux livres de systèmes dynamiques, voir par exemple [KH, §2.5], [AP, 3.5].



Soit  $Q = ]0, 1[^2$  le carré unité (ouvert). Découpons-le en cinq rectangles verticaux de même largeur et notons  $P_0 = ]\frac{1}{5}, \frac{2}{5}[ \times ]0, 1[$  et  $P_1 = ]\frac{3}{5}, \frac{4}{5}[ \times ]0, 1[$  les deuxième et quatrième rectangles verticaux. Effectuons une dilatation horizontale de rapport 5, et une contraction verticale de rapport 1/5, puis replions (de manière lisse) le rectangle ainsi obtenu en un fer à cheval. Faisons en sorte que la composition de ces transformations, notée  $\phi$ , envoie le rectangle vertical  $P_1$  sur le rectangle horizontal  $Q_1 = ]0, 1[ \times ]\frac{1}{5}, \frac{2}{5}[$  par une dilatation horizontale de rapport 5, une contraction verticale de rapport 1/5 et puis une translation, et de même  $P_0$  sur le rectangle horizontal  $Q_0 = ]0, 1[ \times ]\frac{3}{5}, \frac{4}{5}[$ , avec en plus une rotation d'angle  $\pi$ . Le rectangle vertical du milieu  $] \frac{2}{5}, \frac{3}{5}[ \times ]0, 1[$  est envoyé sur la partie courbe du fer à cheval. Nous pouvons faire en sorte que la transformation envoie la réunion

$$R = [0, 1]^2 \cup \left\{ z \in \mathbb{C} : \operatorname{Re} z \leq 0, \left| z - \frac{i}{2} \right| \leq \frac{1}{2} \right\} \cup \left\{ z \in \mathbb{C} : \operatorname{Re} z \geq 1, \left| z - 1 - \frac{i}{2} \right| \leq \frac{1}{2} \right\}$$

du carré fermé et de ses deux demi-disques latéraux dans l'intérieur d'elle-même. Il est possible d'étendre (radialement en dehors de  $R$ ) de manière lisse la transformation  $\phi$  de  $R$  dans lui-même en un difféomorphisme encore noté  $\phi$  de la sphère  $\mathbb{S}_2$ , identifiée à  $\mathbb{R}^2 \cup \{\infty\}$  par la projection stéréographique, ayant un point fixe répulsif en  $\infty$ .



L'intersection  $Q \cap \phi(Q)$  est la réunion  $Q_0 \cup Q_1$  de 2 rectangles horizontaux d'adhérences disjointes  $Q_0$  et  $Q_1$ . L'intersection  $Q \cap \phi(Q) \cap \phi^2(Q)$  est la réunion de 4 rectangles horizontaux d'adhérences deux à deux disjointes

$$Q_{00} = Q_0 \cap \phi(Q_0), \quad Q_{01} = Q_0 \cap \phi(Q_1), \quad Q_{10} = Q_1 \cap \phi(Q_0), \quad Q_{11} = Q_1 \cap \phi(Q_1).$$



Par récurrence, pour tout  $n \geq 1$ , l'intersection  $\bigcap_{i=0}^n \phi^i(Q)$  est réunion de  $2^n$  rectangles horizontaux d'adhérences deux à deux disjointes, chacun d'entre eux de la forme

$$Q_w = \bigcap_{i=0}^{n-1} \phi^i(Q_{w_i})$$

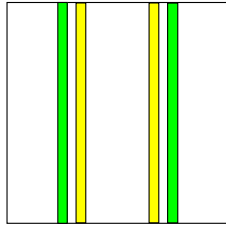
où  $w = (w_0, w_1, \dots, w_{n-1}) \in \{0, 1\}^n$ . Donc  $\Lambda_+ = \bigcap_{i \in \mathbb{N}} \phi^i(Q)$  est un produit  $]0, 1[ \times K_5$  où  $K_5$  est un espace de Cantor dans  $[0, 1]$ . De plus,

$$\Lambda_+ = \bigcup_{\omega = (\omega_i)_{i \in \mathbb{N}} \in \{0, 1\}^{\mathbb{N}}} \bigcap_{i \in \mathbb{N}} \phi^i(Q_{\omega_i}),$$

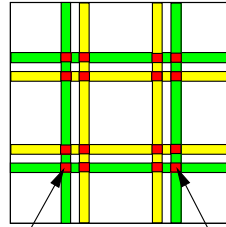
où chaque  $\bigcap_{i \in \mathbb{N}} \phi^i(Q_{\omega_i})$  pour  $\omega = (\omega_i)_{i \in \mathbb{N}} \in \{0, 1\}^{\mathbb{N}}$  est un segment horizontal de la forme  $]0, 1[ \times \{*\omega\}$ .

De même, l'intersection  $\phi^{-1}(Q) \cap Q$ , qui est l'ensemble des points de  $Q$  dont l'image par  $\phi$  est encore dans  $Q$ , est la réunion des 2 rectangles verticaux d'adhérences disjointes  $P_0$  et  $P_1$ . L'intersection  $\phi^{-2}(Q) \cap \phi^{-1}(Q) \cap Q$  est la réunion de 4 rectangles verticaux d'adhérences disjointes

$$P_{00} = \phi^{-1}(P_0) \cap P_0, \quad P_{01} = \phi^{-1}(P_0) \cap P_1, \quad P_{10} = \phi^{-1}(P_1) \cap P_0, \quad P_{11} = \phi^{-1}(P_1) \cap P_1.$$



$$\begin{array}{l} P_{10} = \phi^{-1}(P_1) \cap P_0 \\ P_{00} = \phi^{-1}(P_0) \cap P_0 \end{array}$$



$$\begin{array}{l} P_{11} \cap Q_{11} = \phi^{-1}(P_1) \cap P_1 \cap \phi(P_1) \cap \phi^2(P_1) \\ P_{10} \cap Q_{11} = \phi^{-1}(P_1) \cap P_0 \cap \phi(P_1) \cap \phi^2(P_1) \end{array}$$

Par récurrence, pour tout  $n \geq 1$ , l'intersection  $\bigcap_{i=0}^n \phi^{-i}(Q)$  est réunion de  $2^n$  rectangles verticaux d'adhérences disjointes, chacun d'entre eux de la forme

$$P_w = \bigcap_{i=-n+1}^0 \phi^i(P_{w_i})$$

où  $w = (w_{-n+1}, w_{-n+2}, \dots, w_1, w_0) \in \{0, 1\}^n$ . Donc  $\Lambda_- = \bigcap_{i \in \mathbb{N}} \phi^{-i}(Q)$  est un produit  $K_5 \times ]0, 1[$  où  $K_5$  est un espace de Cantor dans  $[0, 1]$ . De plus,

$$\Lambda_- = \bigcup_{\omega = (\omega_i)_{i \in -\mathbb{N}} \in \{0, 1\}^{-\mathbb{N}}} \bigcap_{i \in -\mathbb{N}} \phi^i(P_{\omega_i}),$$

où chaque  $\bigcap_{i \in -\mathbb{N}} \phi^i(P_{\omega_i})$  pour  $\omega = (\omega_i)_{i \in -\mathbb{N}} \in \{0, 1\}^{-\mathbb{N}}$  est un segment vertical de la forme  $\{*\omega\} \times ]0, 1[$ .

Notons alors

$$\Lambda = \Lambda_- \cap \Lambda_+ = \bigcap_{i \in \mathbb{Z}} \phi^i(Q)$$

appelé *l'attracteur de Smale*. Notons que  $Q$  est un voisinage ouvert de  $\Lambda$ . Par la construction, pour tout  $x \in \mathbb{R}^2$ , nous avons  $\lim_{n \rightarrow +\infty} d(\phi^n(x), \Lambda) = 0$ , et nous dirons que  $\mathbb{R}^2$  est le *bassin d'attraction* de  $\Lambda$  pour  $\phi$  dans  $\mathbb{S}_2 = \mathbb{R}^2 \cup \{\infty\}$ . Le *bassin de répulsion* de  $\infty$  pour  $\phi$ , c'est-à-dire le bassin d'attraction de  $\infty$  pour  $\phi^{-1}$ , c'est-à-dire l'ensemble des  $x \in \mathbb{S}_2$  tels que  $\lim_{n \rightarrow +\infty} \phi^{-n}(x) = \infty$ , est le complémentaire de l'attracteur de Smale : nous avons

$$\{x \in \mathbb{S}_2 : \lim_{n \rightarrow +\infty} \phi^{-n}(x) = \infty\} = \mathbb{S}_2 - \Lambda .$$

En particulier, pour tout voisinage d'adhérence compacte  $U$  de  $\Lambda$ , nous avons (une propriété sur laquelle nous reviendrons plus tard)

$$\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U) = \Lambda . \quad (97)$$

**Proposition 10.14.** *Le fer à cheval  $\phi : \mathbb{S}_2 \rightarrow \mathbb{S}_2$  est un  $C^\infty$ -difféomorphisme ayant exactement deux minimaux,  $\{\infty\}$  qui est un point fixe répulsif, et  $\Lambda = \bigcap_{i \in \mathbb{Z}} \phi^i(Q)$ , qui est une partie hyperbolique pour  $\phi$  homéomorphe à un espace de Cantor. La restriction de  $\phi$  à  $\mathbb{R}^2$  est un difféomorphisme Axiome A de  $\mathbb{R}^2$ , d'ensemble non errant  $\Omega(\phi|_{\mathbb{R}^2}) = \Lambda$ .*

*La restriction de  $\phi$  à  $\Lambda$  est un système dynamique topologique à temps discret, topologiquement mélangeant, minimal, topologiquement conjugué au décalage de Bernoulli bilatère  $(\Sigma = \{0, 1\}^{\mathbb{Z}}, \sigma)$  sur l'alphabet  $\mathcal{A} = \{0, 1\}$ . L'ensemble des points périodiques de  $\phi$  dans  $\Lambda$  est dense dans  $\Lambda$  :*

$$\overline{\text{Per}(\phi|_{\mathbb{R}^2})} = \overline{\text{Per}(\phi|_{\Lambda})} = \Lambda .$$

**Démonstration.** L'espace  $\Lambda$  est le produit  $K_5 \times K_5$  de deux espaces de Cantor, donc est un espace de Cantor, et en particulier compact et non vide.

Rappelons que  $Q_i = \phi(P_i)$  pour  $i = 0, 1$ . Pour tout  $\omega = (\omega_i)_{i \in \mathbb{Z}} \in \Sigma = \{0, 1\}^{\mathbb{Z}}$ , l'intersection  $\bigcap_{i \in \mathbb{Z}} \phi^i(P_{\omega_i})$  est réduite à un singleton, intersection du segment vertical  $\bigcap_{i \in -\mathbb{N}} \phi^i(P_{\omega_i})$  et du segment horizontal  $\bigcap_{i \in \mathbb{N} - \{0\}} \phi^i(P_{\omega_i}) = \bigcap_{i \in \mathbb{N}} \phi^i(Q_{\omega_i})$ , dont nous notons  $h(\omega)$  l'unique point. L'application  $h : \Sigma \rightarrow \Lambda$  ainsi définie est continue et bijective, entre deux espaces compacts, donc est un homéomorphisme. Par construction,  $h$  conjugue le décalage  $\sigma : \Sigma \rightarrow \Sigma$  à la restriction de  $\phi$  à  $\Lambda$ . Par l'exercice E.48, la transformation  $\phi|_{\Lambda}$  est donc non errante, topologiquement mélangeante, minimale, à orbites périodiques denses, comme le décalage de Bernoulli bilatère sur l'alphabet  $\mathcal{A} = \{0, 1\}$ .

La décomposition de  $\mathbb{R}^2$  en produit  $\mathbb{R} \times \mathbb{R}$  de l'axe horizontal par l'axe vertical donne la décomposition cherchée (qui est constante) de l'espace tangent en tout point  $x$  de  $\Lambda$  en  $E_x^u \oplus E_x^s$ , car pour tout  $x \in \Lambda$ , l'automorphisme linéaire  $T_x \phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  est hyperbolique de matrice diagonale  $\begin{pmatrix} 5 & 0 \\ 0 & \frac{1}{5} \end{pmatrix}$ . Donc  $\Lambda$  est une partie hyperbolique de  $\mathbb{S}_2$  pour  $\phi$ .

Puisque  $\Lambda$  est un attracteur de bassin d'attraction  $\mathbb{R}^2$ , l'ensemble non errant de la restriction de  $\phi$  à  $\mathbb{R}^2$  est égal à l'ensemble non errant de sa restriction à  $\Lambda$ , donc égal à  $\Lambda$ . Par conséquent,  $\phi|_{\mathbb{R}^2}$  est un difféomorphisme Axiome A.  $\square$

## 10.6 Le critère des champs de cônes et la stabilité des ensembles hyperboliques

Les champs de cônes sont un outil primordial pour monter l'existence, dans un ouvert d'une variété, de parties hyperboliques pour un difféomorphisme de cette variété.

Soient  $M$  une variété différentielle lisse,  $\phi : M \rightarrow M$  un  $C^1$ -difféomorphisme,  $\Lambda$  une partie de  $M$  invariante par  $\phi$  et  $\| \cdot \| = (\| \cdot \|_x)_{x \in \Lambda}$  une norme (continue) sur le fibré vectoriel  $TM|_{\Lambda}$  vérifiant la propriété (92) (ce qui est automatique si  $\Lambda$  est compact).

Un *champ de cônes* sur  $\Lambda$  est une famille  $(C_x)_{x \in \Lambda}$  de parties de  $TM|_{\Lambda}$  telle qu'il existe  $c > 0$  tel que, pour tout  $x \in \Lambda$ , il existe une décomposition en somme directe vectorielle  $T_x M = E_x^1 \oplus E_x^2$ , et des normes  $\| \cdot \|_{1,x}$  et  $\| \cdot \|_{2,x}$  sur  $E_x^1$  et  $E_x^2$  respectivement telles que

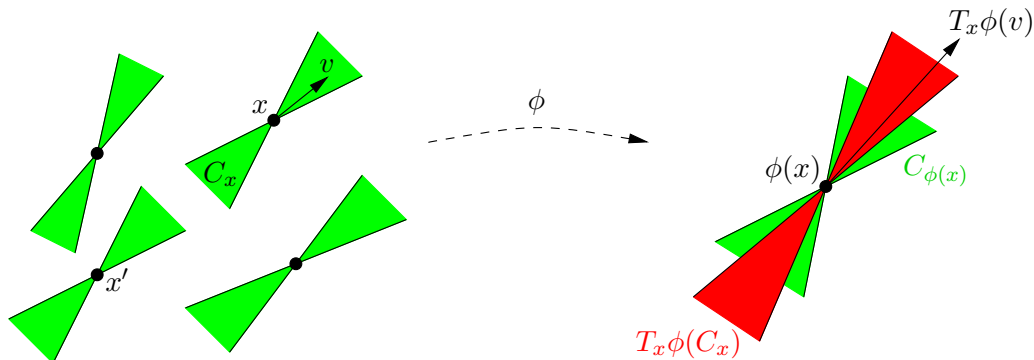
- **(condition de cône)**  $C_x = \{v = v_1 + v_2 : v_1 \in E_x^1, v_2 \in E_x^2, \|v_1\|_{1,x} \leq \|v_2\|_{2,x}\}$ ,
- **(condition de compatibilité avec la norme)** pour tous les  $v_1 \in E_x^1$  et  $v_2 \in E_x^2$ , nous avons  $\frac{1}{c} \|v_1 + v_2\|_x \leq \max\{\|v_1\|_{1,x}, \|v_2\|_{2,x}\} \leq c \|v_1 + v_2\|_x$ .

Remarquons que nous ne demandons pas d'hypothèse de continuité en  $x$ , ni de la décomposition  $T_x M = E_x^1 \oplus E_x^2$ , ni des normes  $\| \cdot \|_{1,x}$  et  $\| \cdot \|_{2,x}$ , ni de  $C_x$ .

Notons que pour tout  $x \in M$ , le cône  $C_x$  détermine les dimensions  $d_x^1 = \dim(E_x^1)$  et  $d_x^2 = \dim(E_x^2)$  des sous-espaces vectoriels  $E_x^1$  et  $E_x^2$  : l'entier  $d_x^1$  est la dimension maximale d'un sous-espace vectoriel de  $T_x M$  ne rencontrant  $C_x$  qu'en  $\{0\}$  et l'entier  $d_x^2$  est la dimension maximale d'un sous-espace vectoriel de  $T_x M$  contenu dans  $C_x$ .

Notons que la famille  $(C_x^* = T_x M - \overset{\circ}{C}_x)_{x \in \Lambda}$  des complémentaires des intérieurs des cônes  $C_x$  pour  $x \in \Lambda$  est aussi un champ de cônes sur  $\Lambda$ , obtenu en échangeant  $E_x^1$  et  $E_x^2$ .

Notons que si  $\Lambda'$  est une partie de  $\Lambda$ , alors  $(C_x)_{x \in \Lambda'}$  est un champ de cônes sur  $\Lambda'$  (pour la restriction de la norme  $\| \cdot \|$  de  $TM|_{\Lambda}$  à  $TM|_{\Lambda'}$ ).



**Proposition 10.15. (Critère des champs de cônes instables)** Une partie compacte et invariante  $\Lambda$  de  $M$  est hyperbolique pour  $\phi$  si et seulement s'il existe un champ de cônes  $(C_x)_{x \in \Lambda}$  sur  $\Lambda$ , des constantes  $\lambda, \mu > 1$  et  $N \in \mathbb{N} - \{0\}$ , tels que, pour tout  $x \in \Lambda$ ,

- **(condition d'inclusion stricte)** nous avons  $\dim(E_x^i) = \dim(E_{\phi(x)}^i)$  pour  $i = 1, 2$  et  $T_x \phi(C_x) \subset \widetilde{C_{\phi(x)}} = \{v = v_1 + v_2 : v_1 \in E_{\phi(x)}^1, v_2 \in E_{\phi(x)}^2, \lambda \|v_1\|_{1,\phi(x)} \leq \|v_2\|_{2,\phi(x)}\}$ ,
- **(condition de dilatation)** pour tous les  $v \in C_x$  et  $v' \in C_x^*$ , nous avons

$$\|T_x \phi^N(v)\|_{\phi^N(x)} \geq \mu \|v\|_x \quad \text{et} \quad \|T_x \phi^{-N}(v')\|_{\phi^{-N}(x)} \geq \mu \|v'\|_x .$$

Notons que  $\widetilde{C}_x \subset C_x$  (et même  $\widetilde{C}_x - \{0\} \subset \overset{\circ}{C}_x$ ) pour tout  $x \in \Lambda$ .

**Démonstration.** Nous traitons d'abord le sens direct, plus facile, de cette équivalence. Le sens réciproque, qui est fort utile pour la construction de parties hyperboliques pour un difféomorphisme, se démontrera par le même principe que pour la démonstration de l'exercice E.60 : pour construire l'espace stable en un point  $x$ , on considère le cône stable en  $\phi^{-n}(x)$ , et on le pousse en avant par  $\phi^n$  ; les deux propriétés de contraction donnée par le critère garantissent alors une convergence du cône image vers l'espace stable cherché.

Supposons tout d'abord que  $\Lambda$  soit une partie hyperbolique pour  $\phi$ . Remplaçons la norme de  $TM|_\Lambda$  par une norme adaptée à  $\phi$  (voir la remarque (3) suivant la démonstration de l'équivalence des définitions d'une partie hyperbolique). Par la compacité de  $\Lambda$ , posons  $\lambda_u = \max_{x \in \Lambda} \|T_x(\phi^{-1})|_{E_x^u}\| < 1$  et  $\lambda_s = \max_{x \in \Lambda} \|T_x\phi|_{E_x^s}\| < 1$ , et prenons

$$E_x^1 = E_x^s, \quad E_x^2 = E_x^u, \quad \| \cdot \|_{1,x} = \| \cdot \|_x |_{E_x^s}, \quad \| \cdot \|_{2,x} = \| \cdot \|_x |_{E_x^u},$$

$$\lambda = \frac{1}{\lambda_u \lambda_s} > 1, \quad \mu = \min\left\{\frac{1}{\lambda_u}, \frac{1}{\lambda_s}\right\} > 1, \quad N = 1.$$

Considérons la famille  $(C_x)_{x \in \Lambda}$  de parties de  $TM|_\Lambda$  définies pour tout  $x \in \Lambda$  par

$$C_x = \{v_1 + v_2 : v_1 \in E_x^1, v_2 \in E_x^2, \|v_1\|_{1,x} \leq \|v_2\|_{2,x}\}.$$

Elle vérifie par construction la **condition de cône**. La **condition de compatibilité avec la norme** pour la famille  $(C_x)_{x \in \Lambda}$  découle du fait que la norme de  $TM|_\Lambda$  est adaptée. Montrons la **condition d'inclusion stricte**. Puisque  $T_x\phi$  est un isomorphisme linéaire préservant les sous-espaces stables et instables, nous avons

$$\dim(E_{\phi(x)}^i) = \dim(T_x\phi(E_x^i)) = \dim(E_x^i).$$

De plus, pour tous les  $x \in \Lambda$  et  $v \in C_x$ , c'est-à-dire  $v = v_u + v_s \in T_xM$  avec  $v_u \in E_x^u$ ,  $v_s \in E_x^s$  et  $\|v_s\|_x \leq \|v_u\|_x$ , nous avons

$$\begin{aligned} \lambda \| (T_x\phi(v))_s \|_{1,\phi(x)} &= \lambda \| (T_x\phi(v))_s \| = \frac{1}{\lambda_u \lambda_s} \| T_x\phi(v_s) \| \leq \frac{1}{\lambda_u} \| v_s \| \\ &\leq \frac{1}{\lambda_u} \| v_u \| \leq \| T_x\phi(v_u) \| = \| (T_x\phi(v))_u \| = \| (T_x\phi(v))_u \|_{2,\phi(x)}. \end{aligned}$$

Ceci montre que  $T_x\phi(C_x) \subset \widetilde{C_{\phi(x)}}$  comme voulu. Montrons enfin la **condition de dilatation**. Soient  $x \in \Lambda$ ,  $v \in C_x$  et  $v' \in C_x^*$ . Puisque  $N = 1$  et  $T_x\phi(C_x) \subset \widetilde{C_{\phi(x)}} \subset C_{\phi(x)}$ , et puisque la norme est adaptée, nous avons

$$\begin{aligned} \| T_x\phi^N(v) \|_{\phi^N(x)} &= \max\{\| T_x\phi(v)_s \|, \| T_x\phi(v)_u \| \} = \| (T_x\phi(v))_u \| = \| T_x\phi(v_u) \| \\ &\geq \frac{1}{\lambda_u} \| v_u \| = \frac{1}{\lambda_u} \| v \| \geq \mu \| v \|_x. \end{aligned}$$

De même, puisque  $T_x\phi^{-1}(C_x^*) \subset C_{\phi^{-1}(x)}^*$ , nous avons

$$\begin{aligned} \| T_x\phi^{-N}(v') \|_{\phi^{-N}(x)} &= \max\{\| T_x\phi^{-1}(v')_s \|, \| T_x\phi^{-1}(v')_u \| \} = \| (T_x\phi^{-1}(v'))_s \| \\ &= \| T_x\phi^{-1}(v'_s) \| \geq \frac{1}{\lambda_s} \| v'_s \| = \frac{1}{\lambda_s} \| v' \| \geq \mu \| v' \|_x. \end{aligned}$$

Réciproquement, soit  $(C_x)_{x \in \Lambda}$  un champ de cônes sur  $\Lambda$  vérifiant les propriétés d'inclusion stricte et de dilatation pour  $\phi$ . Montrons qu'il existe, pour tout  $x \in \Lambda$ , des (uniques) sous-espaces vectoriels  $E_x^u$  et  $E_x^s$  de  $T_x M$  tels que, avec les notations des remarques précédant l'énoncé de la proposition 10.15,

- (i)  $E_x^s \subset C_x^*$  et  $\dim(E_x^s) = d_x^1$
- (ii)  $E_x^u \subset C_x$  et  $\dim(E_x^u) = d_x^2$
- (iii)  $T_x \phi(E_x^u) = E_{\phi(x)}^u$  et  $T_x \phi(E_x^s) = E_{\phi(x)}^s$ .

Les conditions (i) et (ii) montreront en particulier que  $T_x M = E_x^s \oplus E_x^u$ . Avec la **condition d'inclusion stricte**, le point (iii) montrera que

$$E_x^u = T_{\phi^{-1}(x)} \phi(E_{\phi^{-1}(x)}^u) \subset T_{\phi^{-1}(x)} \phi(C_{\phi^{-1}(x)}) \subset \widetilde{C}_x.$$

Par la **condition de compatibilité avec la norme** de  $(C_x)_{x \in \Lambda}$ , ceci montrera qu'en posant  $E^s = \bigcup_{x \in \Lambda} E_x^s$  et  $E^u = \bigcup_{x \in \Lambda} E_x^u$ , nous avons une décomposition en somme directe  $\Gamma_b(TM|_\Lambda) = \Gamma_b(E^s) \oplus \Gamma_b(E^u)$ . Par la **condition de dilatation**, cette décomposition vérifiera l'assertion (2) de la définition 10.11, donc  $\Lambda$  sera une partie hyperbolique de  $M$  pour  $\phi$ .

Maintenant, fixons un point  $x \in \Lambda$ , et discutons suivant la nature de l'orbite (complète)  $\mathcal{O}(x) = \{\phi^n(x) : n \in \mathbb{Z}\}$  de  $x$  par  $\phi$ .

**Cas 1 :** Supposons que  $x$  n'est pas périodique. Pour tout  $y \in \Lambda$ , notons  $G(y)$  le sous-espace fermé donc compact<sup>221</sup> non vide des sous-espaces vectoriels de  $T_y M$  de dimension  $d_y^2$  contenus dans le cône fermé  $C_y$ . L'application linéaire  $T_y \phi$  induit une application continue  $(T_y \phi)_*$  de  $G(y)$  dans  $G(\phi(y))$ , car  $T_y \phi$  envoie  $C_y$  dans  $C_{\phi(y)}$ . En tant qu'intersection décroissante de compacts non vide de  $C_x$ , l'intersection

$$G = \bigcap_{n \in \mathbb{N}} (T_{\phi^{-n}(x)} \phi^n)_* (G(\phi^{-n}(x)))$$

est non vide, en fait réduit à un singleton par la **condition d'inclusion stricte**. Notons  $E_x^u$  son élément, qui vérifie la condition (ii). Notons que pour tout  $n \in \mathbb{Z}$ , le point  $\phi^n(x)$  n'est pas non plus périodique, et que par unicité, nous avons  $T_x \phi(E_x^u) = E_{\phi(x)}^u$ , ce qui montre la première partie de l'assertion (iii). En remplaçant  $\phi$  par  $\phi^{-1}$ , nous construisons  $E_x^s$ , qui vérifie (i) et la seconde partie de l'assertion (iii).

**Cas 2 :** Supposons que  $x$  est périodique, de période  $k \geq 1$ . Notons  $A = T_x \phi^k$ , qui est un automorphisme linéaire de  $T_x M$ . Si  $A$  admettait une valeur propre (complexe) de module 1, alors il existerait  $c > 0$  et un vecteur non nul  $v \in T_x M$  (n'importe quel vecteur non nul dans le sous-espace vectoriel stable correspondant à l'ensemble des valeurs propres complexes de module 1 de  $A$ ) tel que  $\frac{1}{c} \|v\|_x \leq \|A^n v\|_x \leq c \|v\|_x$  pour tout  $n \in \mathbb{Z}$ , ce qui contredirait la **condition de dilatation** du champ de cônes.

Donc  $A$  est un automorphisme linéaire hyperbolique, et nous notons  $E_x^s$  (respectivement  $E_x^u$ ) le sous-espace stable (respectivement instable) de  $T_x M$  pour  $A$ . Par la **condition de dilatation** du champ de cônes, nous avons  $E_x^s \cap C_x = E_x^u \cap C_x^* = \{0\}$ , ce qui démontre les propriétés (i) et (ii). Puisque  $T_{\phi^n(x)} \phi \circ T_x \phi^n = T_x \phi^{n+1} = T_{\phi(x)} \phi^n \circ T_x \phi$ , la propriété (iii) est vérifiée.  $\square$

221. dans la variété grassmannienne  $\mathcal{G}_{d_y}^2(T_y M)$  des sous-espaces vectoriels de dimension  $d_y^2$  de l'espace vectoriel réel  $T_y M$ , voir par exemple [Pau2, §2.4.3]

Ce critère d'hyperbolicité par les champs de cônes permet de montrer la stabilité de l'existence d'une partie hyperbolique par petite perturbation. Nous avons besoin de donner un sens précis à cette notion de "petite perturbation" de système dynamique différentiable.

Soit  $k \in \mathbb{N} - \{0\}$ . Rappelons que toute variété différentielle lisse est supposée métrisable et séparable (donc  $\sigma$ -compacte et à base dénombrable de voisinages, voir par exemple [Pau2, §2.1]), et munie d'une distance induisant sa topologie.

Pour toutes les variétés différentielles lisses  $M, N$  et les applications  $f, g : M \rightarrow N$  de classe  $C^k$ , nous définissons par récurrence les espaces tangents itérés de  $M$  par  $T^0M = M$  et  $T^nM = T(T^{n-1}M)$  si  $n \geq 1$ , et les applications tangentes itérées de  $f$  par  $T^0f = f$  et  $T^n f = T(T^{n-1}f) : T^nM \rightarrow T^nN$  si  $1 \leq n \leq k$ . En notant  $K_i$  un compact de  $T^iM$  pour tout  $i = 0, \dots, k$ , et  $\mathcal{K} = (K_i)_{0 \leq i \leq k}$ , notons

$$d_{\mathcal{K}}(f, g) = \sum_{i=0}^k \sup_{x \in K_i} d(T^i f(x), T^i g(x)).$$

La topologie  $C^k$  (faible) sur l'ensemble  $C^k(M, N)$  des applications  $C^k$  de  $M$  dans  $N$  est la topologie métrisable<sup>222</sup> définie par la famille de pseudo-distances  $(d_{\mathcal{K}})_{\mathcal{K}}$  (voir par exemple [Pau3, §2.2] pour une définition, complètement analogue à celle qui définit la topologie induite par une seule vraie distance).

On montre (voir par exemple [Hir]) que  $\text{Diff}^k(M)$  est ouvert dans  $C^k(M, M)$ , que c'est un espace de Baire (voir l'exercice E.46 pour une définition), que c'est une variété banachique si  $M$  est compacte (mais nous n'aurons pas besoin de ce dernier point, donc nous ne donnons pas de définition), et que la composition des applications  $C^k$  est continue pour la topologie  $C^k$ . La définition convient si  $k = \infty$  en prenant toutes les suites finies de compacts  $\mathcal{K} = (K_i)_{0 \leq i \leq n}$  pour  $n \in \mathbb{N}$ , sauf que la topologie n'est pas forcément métrisable, et la variété  $\text{Diff}^\infty(M)$  n'est pas banachique, mais de Fréchet si  $M$  est compacte.

En pratique, pour travailler avec la topologie  $C^k$ , on regarde la convergence uniforme, dans les compacts des cartes locales, des applications et de toutes leurs dérivées partielles jusqu'à l'ordre  $k$ .

Par exemple, lorsque  $k = 1$ , la topologie  $C^1$  sur le groupe  $\text{Diff}^1(M)$  est définie par la famille de pseudo-distances  $d_{K, K'}$ , où pour tous les compacts  $K$  de  $M$  et  $K'$  de  $TM$ , pour tous les éléments  $f, g \in \text{Diff}^1(M)$ , nous avons

$$d_{K, K'}(f, g) = \sup_{x \in K} d(f(x), g(x)) + \sup_{v \in K'} d(Tf(v), Tg(v)).$$

Lorsque  $M = \mathbb{T}^N$ , nous retrouvons la définition de la topologie  $C^1$  sur  $C^1(\mathbb{T}^N, \mathbb{T}^N)$  donnée avant l'énoncé du théorème 10.8 de stabilité structurelle des difféomorphismes Anosov linéaires du tore.

**Corollaire 10.16. (Théorème de stabilité des parties hyperboliques)** *Soit  $\Lambda$  un compact invariant hyperbolique de  $M$  pour  $\phi$ . Alors il existe un voisinage  $\mathcal{V}$  de  $\phi$  dans*

<sup>222</sup>. Pour  $i = 0, \dots, k$ , soit  $(K_{i,n})_{n \in \mathbb{N}}$  une suite de compacts de  $T^iM$  telle que  $K_{i,n}$  soit contenu dans l'intérieur de  $K_{i,n+1}$  et  $T^iM = \bigcup_{n \in \mathbb{N}} K_{i,n}$ . Alors la distance

$$d(f, g) = \sum_{(n_1, \dots, n_k) \in \mathbb{N}^k} 2^{-(n_1 + \dots + n_k)} \min \{ 1, d_{(K_{1,n_1}, \dots, K_{k,n_k})}(f, g) \}$$

convient.

$\text{Diff}^1(M)$  et un voisinage compact  $U$  de  $\Lambda$  dans  $M$  tels que, pour tout  $\psi \in \mathcal{V}$ , la partie  $\Lambda_\psi = \bigcap_{n \in \mathbb{Z}} \psi^{-n}(U)$  soit un compact invariant hyperbolique (non vide) de  $M$  pour  $\psi$ .

**Démonstration.** Nous pouvons supposer que la norme de  $TM|_\Lambda$  est adaptée pour  $\phi$ . Comme les dimensions de  $E_x^s$  et  $E_x^u$  sont localement constantes en  $x \in \Lambda$ , nous pouvons supposer qu'elles sont constantes, notées  $d^s$  et  $d^u$ . Par partition de l'unité, nous pouvons construire, pour tout  $y$  dans un voisinage compact  $U$  de  $\Lambda$  des sous-espaces vectoriels  $E_y^1$  (respectivement  $E_y^2$ ) de dimension  $d^s$  (respectivement  $d^u$ ) dans  $T_yM$ , dépendant continûment de  $y$  et coïncidant avec  $E_y^s$  (respectivement  $E_y^u$ ) si  $y \in \Lambda$ . Quitte à réduire  $U$ , pour tout  $y \in U$ , nous avons  $T_yM = E_y^1 \oplus E_y^2$ .

De même, quitte à réduire  $U$ , nous étendons continûment la norme de  $TM|_\Lambda$  à  $U$ , de sorte que  $\|v_1 + v_2\|_y = \max\{\|v_1\|_y, \|v_2\|_y\}$  pour tous les  $y \in U$ ,  $v_1 \in E_y^1$  et  $v_2 \in E_y^2$ . Notons  $\|v_1\|_{1,y} = \|v_1\|_y$  et  $\|v_2\|_{2,y} = \|v_2\|_y$  pour tous les  $v_1 \in E_y^1$  et  $v_2 \in E_y^2$ , ce qui définit un champ de cônes  $(C_y)_{y \in U}$  sur le compact  $U$  en posant

$$C_y = \{v_1 + v_2 : v_1 \in E_y^1, v_2 \in E_y^2, \|v_1\|_{1,y} \leq \|v_2\|_{2,y}\}$$

pour tout  $y \in U$ . Notons que les conditions d'inclusion stricte et de dilatation de la proposition 10.15 pour le difféomorphisme  $\phi$  sont vérifiées en tout point  $y \in \Lambda$  puisque la partie  $\Lambda$  est hyperbolique pour  $\phi$  et la norme de  $TM|_\Lambda$  est adaptée pour  $\phi$ .

Pour tout  $\psi \in \text{Diff}^1(M)$ , notons  $\Lambda_\psi = \bigcap_{n \in \mathbb{Z}} \psi^{-n}(U)$  l'ensemble des points de  $U$  dont l'orbite complète par  $\psi$  reste dans  $U$ . Il est fermé dans le compact  $U$ , donc compact. Par continuité et compacité, il existe un voisinage  $\mathcal{V}$  de  $f$  dans  $\text{Diff}^1(M)$  tel que, quitte à restreindre  $U$ , pour tout  $\psi \in \mathcal{V}$ , l'ensemble  $\Lambda_\psi$  est non vide (sinon en prenant des limites, il existerait des points de  $\Lambda$  dont l'orbite par  $\phi$  sortirait de  $\Lambda$ ) et les conditions d'inclusion stricte et de dilatation du champs de cône  $(C_y)_{y \in \Lambda_\psi}$  sur l'ensemble  $\Lambda_\psi$  pour le difféomorphisme  $\psi$  sont encore vérifiées (quitte à remplacer  $\lambda$  et  $\mu$  par  $\lambda' \in ]1, \lambda[$  et  $\mu' \in ]1, \mu[$ ). La partie réciproque de la proposition 10.15 permet de conclure.  $\square$

Le résultat suivant est alors immédiat, puisqu'un difféomorphisme d'Anosov est un difféomorphisme dont la variété toute entière est un ensemble hyperbolique. Notons qu'il existe des variétés différentielles n'admettant pas de difféomorphisme Anosov, donc l'ouvert ci-dessus peut être vide.

**Corollaire 10.17.** *Pour toute variété différentielle lisse  $M$ , l'ensemble des  $C^1$ -difféomorphismes Anosov de  $M$  est un ouvert de  $\text{Diff}^1(M)$  pour la topologie  $C^1$ .*  $\square$

## 10.7 Expansivité et lemme de pistage des systèmes dynamiques hyperboliques

Dans cette partie, nous donnons deux conséquences dynamiques de l'existence d'ensembles hyperboliques de difféomorphismes de variétés.

Soient  $M$  une variété différentielle lisse,  $\pi : TM \rightarrow M$  son fibré tangent,  $\phi : M \rightarrow M$  un  $C^1$ -difféomorphisme, et  $\Lambda$  une partie de  $M$  invariante par  $\phi$ .

Fixons<sup>223</sup> une norme  $\|\cdot\| : TM \rightarrow \mathbb{R}$  (dite *riemannienne*) sur le fibré vectoriel  $TM$  qui est une norme euclidienne sur chaque fibre  $T_xM$  pour  $x \in M$  et dont le carré est  $C^\infty$ .

223. Pour l'existence des objets  $\|\cdot\|$ ,  $\exp$  et  $d$  ci-dessous en général (voir les exemples ci-après), nous renvoyons à un cours de géométrie riemannienne, par exemple [Pau5] :

- la norme  $\|\cdot\| : TM \rightarrow \mathbb{R}$  d'une métrique riemannienne  $g$  lisse convient ;
- l'application  $\exp : v \mapsto (\pi(v), \exp_{\pi(v)}(v))$ , où pour tout  $x \in M$ , l'application  $\exp_x : T_xM \rightarrow M$  est

Fixons un  $C^\infty$ -difféomorphisme  $\exp$  d'un voisinage de l'image de la section nulle<sup>224</sup> de  $TM$  dans un voisinage de la diagonale<sup>225</sup> dans  $M \times M$ , envoyant l'image de section nulle sur la diagonale, tel que le diagramme suivant commute sur le domaine de définition de  $\exp$  :

$$\begin{array}{ccc} TM & \xrightarrow{\exp} & M \times M \\ \pi \searrow & & \swarrow \text{pr}_1 \\ & M & \end{array}$$

Ainsi, pour tous les  $x \in M$  et  $v \in T_x M$  assez petit, nous avons  $\exp(v) = (x, \exp_x v)$  où  $\exp_x$  est un difféomorphisme d'un voisinage ouvert de 0 dans  $T_x M$  sur un voisinage ouvert de  $x$  dans  $M$ , envoyant 0 sur  $x$ . En identifiant l'espace tangent en tout point d'un espace vectoriel (réel de dimension finie)  $E$  avec  $E$ , nous supposons que l'application tangente  $T_0 \exp_x : T_x M \rightarrow T_x M$  de  $\exp_x$  au vecteur nul 0 est l'application identité.

Fixons une distance  $d$  sur  $M$  qui vérifie la propriété suivante de compatibilité avec la norme : pour tout compact  $K$  de  $M$ , il existe  $\epsilon_0 > 0$  et  $c \geq 1$  tels que pour tous les  $x_0, x, y \in K$ , si  $d(x_0, x) \leq \epsilon_0$  et  $d(x_0, y) \leq \epsilon_0$ , alors  $x$  et  $y$  appartiennent au voisinage ouvert de  $x_0$  sur lequel  $\exp_{x_0}$  est un difféomorphisme, et

$$\frac{1}{c} \|\exp_{x_0}^{-1}(y) - \exp_{x_0}^{-1}(x)\| \leq d(x, y) \leq c \|\exp_{x_0}^{-1}(y) - \exp_{x_0}^{-1}(x)\|. \quad (98)$$

**Exemples.** (1) Supposons que  $M$  soit un ouvert  $U$  de l'espace euclidien usuel  $\mathbb{R}^n$  muni de sa norme euclidienne  $\|\cdot\|_{\text{euc}}$  et de sa distance euclidienne  $d_{\text{euc}}$ . Alors  $TM = U \times \mathbb{R}^n$  et  $\pi : TM \rightarrow M$  est la première projection. Nous pouvons prendre pour la norme de  $TM$  l'application  $\|\cdot\|$  définie par

$$\forall (x, v) \in TM = U \times \mathbb{R}^n, \quad \|(x, v)\| = \|v\|_{\text{euc}}.$$

Considérons le voisinage ouvert  $W = \{(x, v) : x \in U, x + v \in U\}$  de l'image de la section nulle  $\{(x, 0) : x \in U\}$  du fibré tangent  $TM = U \times \mathbb{R}^n$ . Nous pouvons prendre pour application  $\exp : W \rightarrow M \times M$  l'application définie par

$$\forall (x, v) \in TM = U \times \mathbb{R}^n, \quad \exp(x, v) = (x, x + v),$$

qui vérifie les propriétés voulues. Pour tout compact  $K$  de  $U$ , soit  $\epsilon_0 = \frac{1}{2} d(K, {}^c U) > 0$ . Alors nous pouvons prendre pour distance  $d$  la distance euclidienne  $d_{\text{euc}}$ , qui vérifie la formule (98) avec  $c = 1$ .

(2) Supposons que  $M$  soit le tore  $\mathbb{T}^N$ . Notons  $\gamma : \mathbb{R}^N \rightarrow \mathbb{T}^N = \mathbb{R}^N / \mathbb{Z}^N$  la projection canonique. Nous identifions, pour tout  $\dot{x} = \gamma(x) \in \mathbb{T}^N$ , l'espace vectoriel euclidien  $\mathbb{R}^N$  à l'espace tangent  $T_{\dot{x}} \mathbb{T}^N$  par l'application  $T_x \gamma : T_x \mathbb{R}^N = \mathbb{R}^N \rightarrow T_{\dot{x}} \mathbb{T}^N$ , de sorte que nous

---

l'application exponentielle en  $x$  de cette métrique riemannienne  $g$ , convient ;

- en notant  $d_g$  la distance riemannienne de cette métrique riemannienne sur toute composante connexe de  $M$  (rappelons que pour tous les  $x, y$  dans une même composante connexe de  $M$ , la distance  $d_g(x, y)$  est la borne inférieure des longueurs riemanniennes des courbes continues de classe  $C^1$  par morceaux entre  $x$  et  $y$ ), nous pouvons prendre la distance  $d$  sur  $M$  définie par  $d(x, y) = \min\{d_g(x, y), 1\}$  si  $x$  et  $y$  sont dans la même composante connexe de  $M$ , et  $d(x, y) = 1$  sinon.

224. c'est-à-dire  $\{0_x : x \in M\}$  où  $0_x$  est le vecteur nul de l'espace vectoriel  $T_x M$

225. c'est-à-dire  $\{(x, x) : x \in M\}$



ayons  $TM = \mathbb{T}^N \times \mathbb{R}^N$  et que  $\pi : TM \rightarrow M$  soit la première projection. Nous pouvons prendre pour la norme de  $TM$  l'application  $\| \cdot \|$  définie par

$$\forall (\dot{x}, v) \in TM = \mathbb{T}^N \times \mathbb{R}^N, \quad \|(\dot{x}, v)\| = \|v\|_{\text{euc}}.$$

Nous pouvons prendre pour application  $\exp$  l'application définie sur tout  $TM$  par

$$\forall (\dot{x}, v) \in TM = \mathbb{T}^N \times \mathbb{R}^N, \quad \exp(\dot{x}, v) = (\dot{x}, \gamma(x + v)).$$

Et nous pouvons prendre pour  $d$  la distance quotient de la distance euclidienne de  $\mathbb{R}^N$ , qui vérifie la formule (98) avec  $\epsilon_0 = \frac{1}{4}$  et  $c = 1$ .

Le principal problème pour parler de proximité pour la topologie  $C^1$  est que deux difféomorphismes proches n'envoient pas forcément une fibre donnée du fibré tangent sur la même fibre du fibré tangent, et qu'il est utile de comparer des fibres proches par un outil global plutôt que de devoir travailler dans chaque trivialisations locale. Le lemme technique suivant permet de contourner ce problème.

**Lemme 10.18. (Procédure de localisation)** *Supposons  $\Lambda$  compact. Il existe  $\theta > 0$  assez petit et un voisinage  $\mathcal{U}$  assez petit de  $\phi$  dans  $\text{Diff}^1(M)$  tels que pour tout  $\psi \in \mathcal{U}$ , il existe une application continue  $\tilde{G} : TM|_{\Lambda} \rightarrow TM|_{\Lambda}$ , que nous noterons  $\tilde{G}_\psi$  en cas de besoin pour marquer sa dépendance en  $\psi$ , telle que  $\pi \circ \tilde{G} = \phi \circ \pi$  et pour tout  $x \in \Lambda$ , l'application  $\tilde{G}_x = \tilde{G}|_{T_x M} : T_x M \rightarrow T_{\phi(x)} M$  vérifie les propriétés suivantes :*

- elle est de classe  $C^1$ ,
- elle est  $C^1$ -proche de l'application linéaire  $T_x \phi : T_x M \rightarrow T_{\phi(x)} M$  uniformément en  $x \in \Lambda$ ,
- elle coïncide avec  $T_x \phi$  sur  $\{v \in T_x M : \|v\|_x \geq 2\theta\}$ , et
- elle coïncide avec  $\exp_{\phi(x)}^{-1} \circ \psi \circ \exp_x$  sur  $\{v \in T_x M : \|v\|_x \leq \theta\}$ .

De plus, si  $\phi$  est de classe  $C^k$  et si  $\psi$  est suffisamment  $C^k$ -proche de  $\phi$ , alors l'application  $\tilde{G}$  s'étend en une application de classe  $C^k$  sur  $TM$  telle que pour tout  $x \in \Lambda$ , l'application  $\tilde{G}_x$  soit  $C^k$ -proche de  $T_x \phi$ . Si  $\psi = \phi$  est de classe  $C^1$ , alors nous avons  $d_0 \tilde{G}_x = T_x \phi$  pour tout  $x \in \Lambda$ .

Il est important de remarquer que dans le diagramme commutatif suivant correspondant à la propriété  $\pi \circ \tilde{G}_\psi = \phi \circ \pi$ , c'est bien la fonction  $\phi$  elle-même et non pas sa perturbation  $\psi$  qui apparaît :

$$\begin{array}{ccc} TM|_{\Lambda} & \xrightarrow{\tilde{G}_\psi} & TM|_{\Lambda} \\ \pi \downarrow & & \downarrow \pi \\ M & \xrightarrow{\phi} & M. \end{array}$$

**Démonstration.** Par la compacité de  $\Lambda$ , il existe  $\theta_0 > 0$  assez petit et  $\mathcal{U}$  un voisinage assez petit de  $\phi$  dans  $\text{Diff}^1(M)$  tels que pour tous les éléments  $\psi \in \mathcal{U}$ ,  $x \in \Lambda$  et  $v \in T_x M$  tels que  $\|v\|_x \leq 2\theta_0$ , le point  $\psi(\exp_x v)$  appartienne à l'image de  $\exp_{\phi(x)}$ . L'application

$$G_x = \exp_{\phi(x)}^{-1} \circ \psi \circ \exp_x$$

est un  $C^1$ -difféomorphisme d'un voisinage ouvert de 0 dans  $T_x M$  dans un voisinage ouvert de 0 dans  $T_{\phi(x)} M$ . Notons que la différentielle en 0 de l'application  $\exp_{\phi(x)}^{-1} \circ \psi \circ \exp_x$

est égale à  $T_x\phi$ , par le théorème de dérivation des fonctions composées et le fait que  $T_0 \exp_x = \text{id}_{T_x M}$ . L'application  $G_x$  est donc  $C^1$ -proche de l'application  $T_x\phi$  sur la partie compacte  $\{v \in T_x M : \|v\|_x \leq 2\theta_0\}$  de manière uniforme en  $x \in \Lambda$ , si  $\theta_0$  et  $\mathcal{U}$  sont assez petits.

Par un argument de recollement standard, nous allons étendre  $G_x$  de manière continue en  $x$  et de classe  $C^1$  sur  $T_x M$  de sorte qu'elle coïncide avec  $T_x\phi$  en dehors d'un voisinage de 0. Pour cela, soit  $\eta : \mathbb{R} \rightarrow \mathbb{R}$  une fonction lisse, valant 1 en  $t \leq 1$  et 0 en  $t \geq 2$ . Pour tout  $0 < \theta < \theta_0$ , posons  $\tilde{G}_x : T_x M \rightarrow T_{\phi(x)} M$  définie par

$$v \mapsto T_x\phi(v) + \eta\left(\frac{\|v\|^2}{\theta^2}\right) (\exp_{\phi(x)}^{-1} \circ \psi \circ \exp_x(v) - T_x\phi(v)) . \quad (99)$$

Considérons l'application  $\tilde{G}$  valant  $\tilde{G}_x$  sur chaque fibre  $T_x M$  de  $TM|_\Lambda$ . La formule (99), et le fait que le carré de la norme riemannienne soit lisse, montrent que si  $\phi$  et  $\psi$  sont de classe  $C^k$  sur  $M$ , alors il existe un petit voisinage ouvert  $U$  de  $\Lambda$  dans  $M$  tel que l'application  $\tilde{G}$  soit définie sur la sous-variété différentielle  $TU \subset TM$  et de classe  $C^k$ . Par partition de l'unité, nous pouvons étendre l'application  $\tilde{G} : TU \rightarrow TM$  sur tout  $TM$  de manière  $C^k$ . Si  $\psi = \phi$  est de classe  $C^1$ , alors l'application  $\tilde{G}_x : T_x M \rightarrow T_{\phi(x)} M$ , qui coïncide avec l'application  $\exp_{\phi(x)}^{-1} \circ \phi \circ \exp_x$  sur un voisinage de 0, a pour différentielle  $d_0 \tilde{G}_x = T_x\phi$  par ce qui précède. Maintenant, si  $\theta_0$  et  $\mathcal{U}$  sont assez petits, l'application  $\tilde{G}$  convient.  $\square$

Le résultat suivant dit en particulier que les orbites par  $\phi$  de deux points distincts d'une partie hyperbolique compacte  $\Lambda$  ne peuvent pas être arbitrairement proches l'une de l'autre pour tous les temps.

Nous dirons qu'un système dynamique topologique  $(X, \psi)$  à temps discret, où  $(X, d)$  est un espace métrique, est (*positivement*) *expansif* s'il existe  $\epsilon > 0$  tel que, si  $x \neq y$ , alors

$$\sup_{n \in \mathbb{N}} d(\psi^n(x), \psi^n(y)) \geq \epsilon,$$

et qu'il est *expansif* si  $\psi$  est inversible, et s'il existe  $\epsilon > 0$  tel que, si  $x \neq y$ , alors

$$\sup_{n \in \mathbb{Z}} d(\psi^n(x), \psi^n(y)) \geq \epsilon .$$

Tout  $\epsilon$  comme ci-dessus est appelé une *constante d'expansivité* de  $\psi$ .

**Exemple.** Un système de Bernoulli  $(\Sigma, \sigma)$  unilatère (respectivement bilatère) sur un alphabet  $\mathcal{A}$  est positivement expansif (respectivement expansif), de constante d'expansivité 1 (pour la distance définie par la formule (4)).

**Proposition 10.19. (Expansivité des systèmes dynamiques hyperboliques)** *Supposons que  $\Lambda$  soit une partie compacte de  $M$  hyperbolique pour  $\phi$ . Il existe un voisinage  $U$  de  $\Lambda$  tel que la restriction de  $\phi$  à  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U)$  soit expansive.*

Notons que l'ensemble  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U)$ , formé des points de  $M$  dont l'orbite complète par  $\phi$  reste dans  $U$ , contient  $\Lambda$  (qui est invariant par  $\phi$ ). Donc ce résultat montre en particulier que le  $C^1$ -difféomorphisme  $\phi$  est expansif sur toute partie hyperbolique compacte. Si  $M = \mathbb{R}^2$ , si  $\phi$  est l'application du fer à cheval de Smale (voir la partie 10.5), si  $\Lambda$  est l'attracteur de Smale, alors pour tout voisinage ouvert  $U$  de  $\Lambda$  dans  $\mathbb{R}^2$ , nous avons  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U) = \Lambda$  (précisément parce que  $\Lambda$  est un attracteur de bassin d'attraction  $\mathbb{R}^2$ ). En général,

$\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U)$  peut être strictement plus grand que  $\Lambda$ . Nous reviendrons dans la partie 10.9 sur ce point.

**Démonstration.** Soit  $U$  un voisinage compact de  $\Lambda$  de sorte que  $\Lambda' = \bigcap_{n \in \mathbb{Z}} \phi^{-n}(U)$  soit un compact hyperbolique pour  $\phi$ , ce qui est possible par le corollaire de stabilité 10.16.

Notons  $\tilde{F} = \tilde{G}_\phi : TM|_{\Lambda'} \rightarrow TM|_{\Lambda'}$  l'application donnée par le lemme 10.18 en prenant  $\psi = \phi$  et  $\Lambda = \Lambda'$ . Si la constante  $\theta$  apparaissant dans ce lemme est assez petite, alors pour tout  $x \in \Lambda'$ , l'application  $\tilde{F}_x - T_x\phi : T_xM \rightarrow T_{\phi(x)}M$  est  $\delta$ -lipschitzienne avec  $\delta$  arbitrairement petit.

L'application  $\tilde{F}$  induit un endomorphisme linéaire borné  $F$  de l'espace de Banach  $\Gamma_b(TM|_{\Lambda'})$ , qui envoie un champ de vecteurs borné pas forcément continu  $\sigma \in \Gamma_b(TM|_{\Lambda'})$  sur le champ de vecteurs borné pas forcément continu

$$F(\sigma) : x \in \Lambda' \mapsto \tilde{F}_{\phi^{-1}(x)}(\sigma(\phi^{-1}(x))) \in T_xM .$$

Si  $T$  est l'endomorphisme  $\phi_*$  défini par la formule (95), la différence  $f = F - T$  est bornée et  $\delta$ -lipschitzienne pour la norme uniforme sur  $\Gamma_b(TM|_{\Lambda'})$ , avec  $\delta$  arbitrairement petit. Par la proposition 10.3, les endomorphismes linéaires continus  $T$  et  $F$  sont conjugués.

Soient  $x$  et  $y$  deux points distincts de  $\Lambda'$  suffisamment proches. Alors il existe un vecteur  $v \in T_xM$  non nul tel que  $y = \exp_x v$ . Soit  $\sigma$  le champ de vecteurs borné (non continu) sur  $\Lambda$ , valant 0 en tout point différent de  $x$  et valant  $v$  en  $x$ . Puisque son orbite complète par  $T$  est non bornée (car  $T$  est un automorphisme linéaire hyperbolique, donc le seul point d'orbite bornée est le vecteur nul), il en est de même de son orbite par  $F$ . Comme  $\tilde{F}_x = \exp_{f(x)}^{-1} \circ \phi \circ \exp_x$  au voisinage de la section nulle par le lemme de localisation 10.18, ceci implique que l'orbite complète de  $y$  ne peut pas rester proche de celle de  $x$ .  $\square$

Pour tout  $\epsilon > 0$ , une  $\epsilon$ -pseudo-orbite de  $\phi$  est une suite  $(x_n)_{n \in \mathbb{Z}}$  dans  $M$  telle que

$$\sup_{n \in \mathbb{Z}} d(x_{n+1}, \phi(x_n)) \leq \epsilon,$$

Le résultat suivant, qui généralise l'exercice E.60 et qui se démontre par des méthodes analogues, dit que les pseudo-orbites sont proches de vraies orbites.

**Théorème 10.20. (Lemme de pistage des systèmes dynamiques hyperboliques)**

*Supposons que  $\Lambda$  soit une partie compacte de  $M$  hyperbolique pour  $\phi$ . Il existe  $\epsilon_0 > 0$ ,  $c_0 \geq 1$  et un voisinage  $U$  de  $\Lambda$  tel que pour tout  $\epsilon \in ]0, \epsilon_0]$ , pour toute  $\epsilon$ -pseudo-orbite  $(x_n)_{n \in \mathbb{Z}}$  de  $\phi$  contenue dans  $U$ , il existe un et un seul  $x \in M$  tel que*

$$\sup_{n \in \mathbb{Z}} d(x_n, \phi^n(x)) \leq c_0 \epsilon .$$

*De plus, si  $\epsilon_0$  et  $U$  sont assez petits, alors le point  $x$  appartient à un compact hyperbolique invariant  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(V)$  où  $V$  est un voisinage compact de  $\Lambda$ .*

Le point  $x$  n'appartient pas forcément à  $\Lambda$ , mais il appartient bel et bien à  $\Lambda$  si celui-ci est de plus supposé être localement maximal (voir la partie 10.9).

**Démonstration.** (Voir par exemple [Yoc2, §2.7].) Soient  $\epsilon_0 > 0$  et  $U$  un voisinage compact de  $\Lambda$ , assez petits (à préciser ultérieurement). Soit  $(x_n)_{n \in \mathbb{Z}}$  une  $\epsilon$ -pseudo-orbite de  $\phi$  dans  $U$  avec  $\epsilon \leq \epsilon_0$ . Pour tout  $i \in \mathbb{Z}$ , soit  $y_i$  l'un des points du fermé  $\Lambda$  les plus proches de  $x_i$ . Si  $\epsilon_0 > 0$  et  $U$  sont assez petits, alors  $x_i$  est proche de  $y_i$ , et par continuité uniforme  $\phi(y_{i-1})$

est assez proche de  $\phi(x_{i-1})$  donc assez proche de  $y_i$ , pour que  $x_i$  soit dans l'image par  $\exp_{y_i}$  d'un voisinage ouvert de 0 sur lequel  $\exp_{y_i}$  soit un difféomorphisme sur un voisinage ouvert de  $y_i$ , et de même  $y_i$  soit dans l'image par difféomorphisme local de  $\exp_{\phi(y_{i-1})}$  pour tout  $i \in \mathbb{Z}$ .

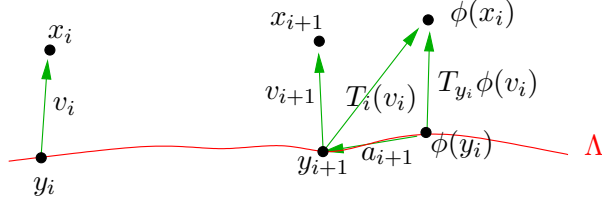
Pour tout  $i \in \mathbb{Z}$ , soient  $E_i = T_{y_i}M$ ,  $E_i^s = E_{y_i}^s$ ,  $E_i^u = E_{y_i}^u$ ,  $v_i \in E_i$  et  $a_i \in T_{\phi(y_{i-1})}M$  tels que

$$\exp_{y_i}(v_i) = x_i \quad \text{et} \quad \exp_{\phi(y_{i-1})}(a_i) = y_i .$$

Notons que  $\dim E_i^s = \dim E_{i+1}^s$  et  $\dim E_i^u = \dim E_{i+1}^u$  si  $\epsilon_0$  et  $U$  sont assez petits.

Pour tout  $i \in \mathbb{Z}$ , soit  $T_i : E_i \rightarrow E_{i+1}$  un isomorphisme linéaire proche de l'application  $v \mapsto -a_{i+1} + T_{y_i}\phi(v)$  (qui va de  $E_i$  dans  $T_{\phi(y_i)}M$ ) tel que

$$T_i(E_i^s) = E_{i+1}^s \quad \text{et} \quad T_i(E_i^u) = E_{i+1}^u .$$



Notons  $\mathcal{E}$  l'espace de Banach

$$\mathcal{E} = \{w = (w_i)_{i \in \mathbb{Z}} \in \prod_{i \in \mathbb{Z}} E_i : \|w\| = \sup_{i \in \mathbb{Z}} \|w_i\| < +\infty\} .$$

Alors  $T : (w_i)_{i \in \mathbb{Z}} \mapsto (T_{i-1}(w_{i-1}))_{i \in \mathbb{Z}}$  est un automorphisme linéaire continu de  $\mathcal{E}$ , qui est hyperbolique si  $\epsilon_0$  et  $U$  sont assez petits, par les propriétés de contraction/dilatation des  $T_{y_i}\phi$ , de sous-espace stable  $\mathcal{E}^s = \prod_{i \in \mathbb{Z}} E_i^s$  et sous-espace instable  $\mathcal{E}^u = \prod_{i \in \mathbb{Z}} E_i^u$ .

Comme dans la démonstration du lemme 10.18, pour tout  $i \in \mathbb{Z}$ , soit  $\tilde{F}_i : E_i \rightarrow E_{i+1}$  tel que si  $\theta > 0$  est assez petit, alors

$$\tilde{F}_i = \begin{cases} T_i & \text{sur } \{v \in E_i : \|v\| \geq 2\theta\} \\ \exp_{y_{i+1}}^{-1} \circ \phi \circ \exp_{y_i} & \text{sur } \{v \in E_i : \|v\| \leq \theta\} , \end{cases} \quad (100)$$

et l'application  $\tilde{F}_i - T_i$  soit  $\delta$ -lipschitzienne pour un  $\delta > 0$  assez petit (indépendant de  $i$ ). Si  $\epsilon_0$  et  $U$  sont assez petits, il existe  $c_1 \geq 1$  tel que pour tout  $i \in \mathbb{Z}$ , nous pouvons supposer que dans la construction des  $v_i$ , nous avons

$$\|v_i\| \leq \frac{\theta}{2} \quad \text{et} \quad \|\tilde{F}_i(v_i) - v_{i+1}\| \leq c_1 \epsilon : \quad (101)$$

ceci est possible car nous avons  $v_{i+1} = \exp_{y_{i+1}}^{-1}(x_{i+1})$ ,  $d(x_{i+1}, \phi(x_i)) \leq \epsilon$  et

$$\tilde{F}_i(v_i) = \exp_{y_{i+1}}^{-1} \circ \phi \circ \exp_{y_i}(v_i) = \exp_{y_{i+1}}^{-1}(\phi(x_i)) . \quad (102)$$

Notons  $F : (w_i)_{i \in \mathbb{Z}} \mapsto (\tilde{F}_{i-1}(w_{i-1}))_{i \in \mathbb{Z}}$ , qui est une application de  $\mathcal{E}$  dans  $\mathcal{E}$  telle que l'application  $f = F - T$  soit bornée (car  $F$  coïncide avec  $T$  en dehors d'un borné) et  $\delta$ -lipschitzienne avec  $\delta > 0$  assez petit. Alors par la proposition 10.3 et la remarque (1) qui suit son énoncé, l'application  $F$  admet un unique point fixe  $w = (w_i)_{i \in \mathbb{Z}}$ . Donc

$$\forall i \in \mathbb{Z}, \quad \tilde{F}_i(w_i) = w_{i+1} .$$

Par la formule (101) de droite, si  $v = (v_i)_{i \in \mathbb{Z}}$ , alors  $\|F(v) - v\| \leq c_1 \epsilon$  par la définition de la norme de  $\mathcal{E}$ . Par la proposition 10.5, il existe donc  $c_2 \geq 1$  tel que

$$\|w - v\| \leq c_2 \|F(v) - v\| \leq c_2 c_1 \epsilon .$$

Posons  $x = \exp_{y_0}(w_0)$ . Alors si  $\epsilon_0$  est assez petit, pour tout  $i \in \mathbb{N}$ , par les formules (100) et (102), nous avons

$$\begin{aligned} \phi^i(x) &= (\exp_{y_i} \circ \tilde{F}_{i-1} \circ \exp_{y_{i-1}}^{-1}) \circ \cdots \circ (\exp_{y_i} \circ \tilde{F}_0 \circ \exp_{y_0}^{-1})(\exp_{y_0}(w_0)) \\ &= \exp_{y_i}(\tilde{F}_{i-1} \circ \cdots \circ \tilde{F}_0(w_0)) = \exp_{y_i}(w_i) . \end{aligned}$$

De même,  $\phi^i(x) = \exp_{y_i}(w_i)$  pour tout  $i \in -\mathbb{N}$ . Puisque  $x_i = \exp_{y_i}(v_i)$  et par la formule (98), il existe donc  $c_3 \geq 1$  tel que

$$d(\phi^i(x), x_i) \leq c_3 \|w_i - v_i\| \leq c_3 c_2 c_1 \epsilon .$$

Donc la pseudo-orbite  $(x_i)_{i \in \mathbb{Z}}$  est à distance comme souhaitée de l'orbite du point  $x$ .

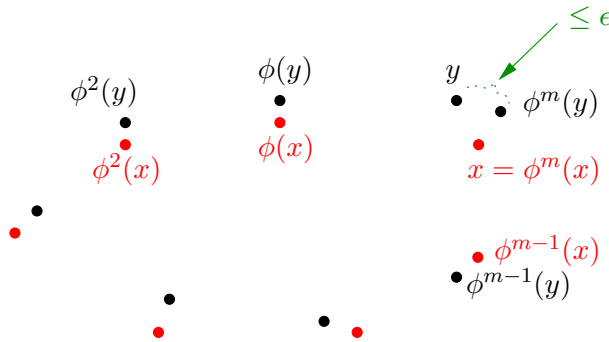
L'unicité de  $x$  découle de la propriété d'expansivité (Proposition 10.19), si  $\epsilon_0$  et  $U$  sont assez petits.

Montrons la dernière affirmation du théorème 10.20. Par le corollaire 10.16, soit  $V$  un voisinage compact de  $\Lambda$  tel que  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(V)$  soit un compact hyperbolique pour  $\phi$ . Si  $\epsilon_0$  et  $U$  sont assez petits, alors le  $(c_0 \epsilon)$ -voisinage fermé de  $U$  est contenu dans  $V$ . Donc pour tout  $n \in \mathbb{Z}$ , le point  $\phi^n(x)$ , qui est à distance au plus  $c_0 \epsilon$  d'un point de  $U$ , appartient à  $V$ . Donc  $x \in \bigcap_{n \in \mathbb{Z}} \phi^{-n}(V)$ .  $\square$

Le corollaire suivant dit qu'une orbite qui se referme presque est à petite distance d'une orbite périodique.

**Corollaire 10.21. (Lemme de fermeture d'Anosov)** *Supposons que  $\Lambda$  soit une partie compacte de  $M$  hyperbolique pour  $\phi$ . Il existe  $\epsilon_0 > 0$ ,  $c_0 \geq 1$ , et un voisinage  $U$  de  $\Lambda$  tel que pour tout  $\epsilon \in ]0, \epsilon_0]$ , pour tout  $m \in \mathbb{N} - \{0\}$  et tout  $y \in U$  tel que  $\phi^k(y) \in U$  pour  $k = 0, \dots, m$  et  $d(\phi^m(y), y) \leq \epsilon$ , il existe un et un seul point  $x \in M$  périodique (de période divisant  $m$ ) tel que*

$$\sup_{0 \leq k \leq m} d(\phi^k(x), \phi^k(y)) \leq c_0 \epsilon .$$



**Démonstration.** Soient  $\epsilon_0, c_0, U$  comme dans l'énoncé du théorème 10.20 et  $\epsilon, m, y$  comme dans l'énoncé du corollaire 10.21. La suite périodique dans  $U$  définie par  $x_{k+pm} = \phi^k(y)$  pour tous les  $k = 0, \dots, m-1$  et  $p \in \mathbb{Z}$  est une  $\epsilon$ -pseudo-orbite de  $\phi$  contenue dans  $U$

par les hypothèses du corollaire 10.21. Donc la suite  $(x_i)_{i \in \mathbb{Z}}$  est à distance au plus  $c_0 \epsilon$  de l'orbite d'un unique point  $x \in M$  par le lemme de pistage (théorème 10.20). Comme c'est aussi le cas pour l'orbite de  $\phi^m(x)$ , nous avons donc  $\phi^m(x) = x$  par l'affirmation d'unicité dans le théorème 10.20.  $\square$

Ce lemme de fermeture permet de montrer qu'un système dynamique hyperbolique a de nombreuses orbites périodiques.

**Corollaire 10.22.** *Supposons que  $\Lambda$  soit une partie compacte de  $M$  hyperbolique pour  $\phi$ , et soit  $V$  un voisinage compact assez petit de  $\Lambda$ . Alors l'ensemble  $\text{Per}(\phi|_{\Lambda'})$  des points périodiques de  $\phi$  dans  $\Lambda' = \bigcap_{n \in \mathbb{Z}} \phi^{-n}(V)$  est dense dans l'ensemble non errant  $\Omega(\phi|_{\Lambda'})$  de la restriction de  $\phi$  à  $\Lambda'$ .*

*En particulier, l'ensemble des points périodiques d'un difféomorphisme Anosov d'une variété compacte est dense dans son ensemble non errant, et un difféomorphisme Anosov d'une variété compacte est Axiome A.*

**Démonstration.** Par le théorème de stabilité des parties hyperboliques (corollaire 10.16), si  $V$  est assez petit, le compact  $\Lambda'$  est une partie hyperbolique pour  $\phi$ . Pour tout  $\epsilon > 0$ , pour tout  $z \in \Omega(\phi|_{\Lambda'})$ , notons  $B$  la boule ouverte de centre  $z$  et de rayon  $\frac{\epsilon}{2c_0+1}$  dans  $\Omega(\phi|_{\Lambda'})$ , où  $c_0 > 0$  est comme dans le lemme de fermeture d'Anosov (corollaire 10.21) pour  $\Lambda'$ . Puisque  $z$  est non errant, il existe  $N \in \mathbb{N}$  tel que l'intersection  $\phi^{-N}(B) \cap B$  soit non vide. Si  $y$  est dans cette intersection, alors  $d(y, \phi^N(y)) < \frac{2\epsilon}{2c_0+1}$  par l'inégalité triangulaire. Donc par le lemme de fermeture d'Anosov (corollaire 10.21), si  $\epsilon$  est assez petit, il existe  $x \in \Lambda'$  tel que  $\phi^N(x) = x$  (donc  $x$  est périodique) et  $d(x, y) \leq \frac{2c_0\epsilon}{2c_0+1}$ . De nouveau par l'inégalité triangulaire, nous avons  $d(x, z) \leq \epsilon$ , donc  $\text{Per}(\phi|_{\Lambda'})$  est dense dans  $\Omega(\phi|_{\Lambda'})$ .  $\square$

Les énoncés 10.20, 10.21 et 10.22 ont des analogues en temps continu, par exemple pour les flots d'Anosov des variétés compactes (voir la fin de la partie 18.1 de [KH] pour une variante plus générale du résultat suivant, et pour une démonstration).

Soient  $(\phi^t)_{t \in \mathbb{R}}$  un flot lisse<sup>226</sup> sur  $M$  (muni de sa norme  $\|\cdot\|$  et de sa distance  $d$  choisies au début de la partie 10.7), et  $\epsilon, \delta > 0$ . Une  $\epsilon$ -pseudo-orbite de  $(\phi^t)_{t \in \mathbb{R}}$  est une courbe lisse  $c : \mathbb{R} \rightarrow M$  telle que pour tout  $t \in \mathbb{R}$ , nous ayons

$$\left\| \dot{c}(t) - \frac{d}{d\tau} \Big|_{\tau=0} \phi^\tau(c(t)) \right\|_{c(t)} \leq \epsilon.$$

Une courbe lisse  $c : \mathbb{R} \rightarrow M$  piste à  $\delta$  près l'orbite par le flot  $(\phi^t)_{t \in \mathbb{R}}$  d'un point  $x \in M$  s'il existe un reparamétrage  $\tau : \mathbb{R} \rightarrow \mathbb{R}$  de classe  $C^1$  de  $c$  tel que  $|\dot{\tau} - 1| \leq \delta$  et pour tout  $t \in \mathbb{R}$ , nous ayons

$$d(c \circ \tau(t), \phi^t(x)) \leq \delta.$$

**Théorème 10.23.** *Si  $M$  est compacte et si  $(\phi^t)_{t \in \mathbb{R}}$  est un flot d'Anosov lisse sur  $M$ , alors pour tout  $\delta > 0$ , il existe  $\epsilon > 0$  tel que toute  $\epsilon$ -pseudo-orbite (respectivement  $\epsilon$ -pseudo-orbite fermée) de  $(\phi^t)_{t \in \mathbb{R}}$  est pistée à  $\delta$  près par une orbite (respectivement orbite périodique) de  $(\phi^t)_{t \in \mathbb{R}}$ .  $\square$*

<sup>226.</sup> Ceci signifie que l'application  $\phi : \mathbb{R} \times M \rightarrow M$  définie par  $(t, x) \mapsto \phi^t(x)$  est de classe  $C^\infty$  (outre que  $\phi^0 = \text{id}$  et  $\phi^{t+\tau} = \phi^t \circ \phi^\tau$  pour tous les  $t, \tau \in \mathbb{R}$ ).

## 10.8 Le théorème des variétés stables et instables

Un problème standard en géométrie différentielle est celui de l'intégrabilité des champs de plans.<sup>227</sup> Étant donné, en tout point  $x$  d'une partie  $\Lambda$  d'une variété différentielle lisse  $M$ , un sous-espace vectoriel  $\Delta_x$  de l'espace tangent  $T_x M$ , il s'agit de trouver une famille de sous-variétés<sup>228</sup> de  $M$  deux à deux disjointes, dont la réunion contient  $\Lambda$ , et tel que l'espace tangent en tout point  $x$  de  $\Lambda$  à la sous-variété passant par  $x$  soit exactement  $\Delta_x$ . Lorsque  $\Lambda = M$  et le champ de plans est lisse et de dimension 1, la situation se ramène localement au problème bien connu d'intégrabilité des champs de vecteurs, auquel cas de telles sous-variétés existent localement, ce sont les courbes intégrales locales de ce champ de vecteurs (voir par exemple [Laf, Pau2]). En dimension plus grande, les situations diffèrent, et donnent lieu à la fois à la théorie des feuilletages (voir par exemple [God2, CanC1, CanC2]) dans le cas intégrable et à celle des champs de contact (voir par exemple [Gei]) dans le cas (fortement) non intégrable.

Dans cette partie, nous étudions le problème de l'intégrabilité des champs de plans stable  $(E_x^s)_{x \in \Lambda}$  et instable  $(E_x^u)_{x \in \Lambda}$  pour une partie hyperbolique  $\Lambda$  d'un  $C^k$ -difféomorphisme. Les méthodes sont encore des méthodes à base du théorème du point fixe d'applications contractantes, et nous commençons par donner la version linéaire à laquelle se ramènera le problème différentiel. Les variétés intégrant la distribution stable seront localement exprimées comme des graphes d'applications de  $E_x^s$  dans  $E_x^u$ , et seront obtenues comme point fixe d'une application de transformation de graphes. Nous suivons de manière proche l'exposition de [Yoc2].

### 10.8.1 Variétés stables des perturbations d'endomorphismes linéaires

Soient  $E$  un espace de Banach réel et  $T$  un endomorphisme linéaire continu hyperbolique de  $E$ , avec  $E = E^s \oplus E^u$  la décomposition associée en sous-espaces stable et instable. Supposons la norme de  $E$  adaptée à  $T$ , et rappelons que la constante d'hyperbolicité de  $T$  est alors  $\text{ch}(T) = \max\{\|T|_{E^s}\|, \|T|_{E^u}^{-1}\|\} < 1$ .

Pour toute application  $F : E \rightarrow E$ , notons

$$W^s(F) = \{x \in E : \sup_{n \in \mathbb{N}} \|F^n(x)\| < \infty\},$$

appelé l'*ensemble stable* de  $F$ . Nous avons

$$F(W^s(F)) \subset W^s(F).$$

Remarquons que  $W^s(T) = E^s$ , qui est un sous-espace vectoriel, donc  $W^s(T)$  est le graphe de l'application nulle de  $E^s$  dans  $E^u$ . Notons que l'endomorphisme  $T$  contracte sur  $W^s(T)$ . Le but du résultat suivant est de montrer que si  $F$  est assez proche de  $T$ , alors l'ensemble stable de  $F$  reste un graphe d'une application proche d'être nulle de  $E^s$  dans  $E^u$ , et de montrer que l'application  $F$  contracte sur son ensemble stable.

**Théorème 10.24.** *Soit  $F = T + f$  où  $f : E \rightarrow E$  est une application  $\delta$ -lipschitzienne avec  $\delta < 1 - \text{ch}(T)$  telle que  $f(0) = 0$ . Alors il existe une application  $\delta'$ -lipschitzienne*

<sup>227</sup>. Notons que cette terminologie, bien qu'usuelle, n'est pas forcément rigoureuse car ici les « plans » peuvent être des sous-espaces vectoriels de n'importe quelle dimension.

<sup>228</sup>. si possible, ou de sous-variétés immergées si nécessaire

$g : E^s \rightarrow E^u$ , avec  $\delta' < 1$  et  $g(0) = 0$ , telle que l'ensemble stable  $W^s(F)$  de  $F$  soit égal au graphe de l'application  $g$ . De plus, nous avons

$$W^s(F) = \{x \in E : \lim_{n \rightarrow +\infty} \|F^n(x)\| = 0\}.$$

En outre, il existe une constante  $\lambda < 1$  telle que  $\|F(x) - F(y)\| \leq \lambda \|x - y\|$  pour tous les  $x, y \in W^s(F)$ .

Pour des raisons de régularité, nous aurons besoin du résultat plus technique suivant, qui implique le précédent en prenant  $\kappa = 1$  dans l'énoncé 10.25, ce qui est possible car  $1 - \text{ch}(T) < \text{ch}(T)^{-1} - 1$ , et en utilisant alors comme constante  $\lambda$  dans l'énoncé 10.24 la valeur

$$\lambda = \delta + \text{ch}(T) < 1.$$

Appelons ensemble  $\kappa$ -stable d'une application  $F : E \rightarrow E$  la partie

$$W_\kappa^s(F) = \{x \in E : \sup_{n \in \mathbb{N}} \kappa^{-n} \|F^n(x)\| < \infty\},$$

**Théorème 10.25.** Soit  $\kappa \in ]\text{ch}(T), 1]$ . Soit  $F = T + f$  où  $f : E \rightarrow E$  est une application  $\delta$ -lipschitzienne avec  $\delta < \min\{\kappa - \text{ch}(T), \text{ch}(T)^{-1} - \kappa\}$  telle que  $f(0) = 0$ . Alors il existe une application  $\delta'$ -lipschitzienne  $G_\kappa : E^s \rightarrow E^u$ , avec  $\delta' < 1$  et  $G_\kappa(0) = 0$ , telle que l'ensemble  $\kappa$ -stable de  $F$  soit le graphe de l'application  $G_\kappa$ . De plus,

$$\forall x, y \in W_\kappa^s(F), \quad \|F(x) - F(y)\| \leq (\delta + \text{ch}(T)) \|x - y\|, \quad (103)$$

et

$$W_\kappa^s(F) = \{x \in E : \lim_{n \rightarrow +\infty} \kappa^{-n} \|F^n(x)\| = 0\}. \quad (104)$$

En outre, si  $\kappa < 1$  et s'il existe  $k \geq 1$  tel que  $f$  soit de classe  $C^k$ , alors l'application  $G_\kappa$  est de classe  $C^k$ , et nous avons  $d_0 G_\kappa = 0$  si  $d_0 f = 0$ .

**Démonstration.** Nous allons utiliser une version à paramètre du théorème du point fixe 8.10 dans la note de bas de page 162, qui se démontre de la même manière. <sup>229</sup>

**Théorème 10.26.** Soient  $X, Z$  deux espaces métriques, avec  $Z$  complet non vide,  $\delta' < 1$  et  $\Theta : X \times Z \rightarrow Z$  une application  $\delta'$ -lipschitzienne, en munissant  $X \times Z$  de la distance  $d((x, z), (x', z')) = \max\{d(x, x'), d(z, z')\}$ . Alors pour tout  $x \in X$ , l'application  $z \mapsto \Theta(x, z)$  admet un et un seul point fixe  $\varphi(x)$  dans  $Z$ , et l'application  $\varphi : X \rightarrow Z$  ainsi définie est  $\delta'$ -lipschitzienne.

Si  $Z$  est un espace de Banach, si  $X$  est un ouvert d'un espace de Banach, et si  $\Theta$  est de classe  $C^k$  pour  $k \in (\mathbb{N} - \{0\}) \cup \{\infty\}$ , alors  $\varphi$  est de classe  $C^k$ , et, en notant respectivement  $d^1$  et  $d^2$  les différentielles par rapport aux première et seconde variables des fonctions définies sur  $X \times Z$ , nous avons

$$d_x \varphi = (\text{id} - d_{(x, \varphi(x))}^2 \Theta)^{-1} \circ d_{(x, \varphi(x))}^1 \Theta. \quad \square \quad (105)$$

<sup>229</sup> Pour tout  $x \in X$ , l'existence de  $\varphi(x)$  est obtenue par application du théorème du point fixe 8.10 dans la note de bas de page 162 à l'application de  $Z$  dans  $Z$  définie par  $z \mapsto \Theta(x, z)$ . Le fait que  $\varphi$  est  $\delta'$ -lipschitzienne s'obtient en prenant  $z = \varphi(x)$  et  $z' = \varphi(x')$  dans la formule  $d(\Theta(x, z), \Theta(x', z')) \leq \delta' \max\{d(x, x'), d(z, z')\}$ . La différentiabilité s'obtient en différentiant l'équation  $\Theta(x, \varphi(x)) = \varphi(x)$ , sachant que puisque  $\Theta$  est  $\delta'$ -lipschitzienne, la norme de l'opérateur linéaire continu  $d_{(x, \varphi(x))}^2 \Theta$  de différentielle en  $(x, \varphi(x))$  de  $\Theta$  par rapport à la seconde variable est au plus  $\delta' < 1$ , donc l'endomorphisme  $\text{id} - d_{(x, \varphi(x))}^2 \Theta$  de  $Z$  est inversible.



Reprenons les notations de la proposition 10.3. Pour toute application  $\gamma : E \rightarrow E$ , notons  $\gamma_s : E \rightarrow E^s$  et  $\gamma_u : E \rightarrow E^u$  les applications telles que  $\gamma = \gamma_s + \gamma_u$ . Notons  $S$  et  $U$  les restrictions de  $T$  à  $E^s$  et  $E^u$  respectivement, de sorte que  $\text{ch}(T) = \max\{\|S\|, \|T^{-1}\|\}$ . Posons  $\mathbb{N}^* = \mathbb{N} - \{0\}$ . Pour tout  $\kappa \in ]\text{ch}(T), 1[$ , par l'hypothèse sur  $\delta$ , nous avons que

$$\delta' = \max\{\kappa^{-1}(\|S\| + \delta), \|U^{-1}\|(\kappa + \delta)\} < 1.$$

Considérons les espaces de Banach

$$\mathcal{E}^s = \left\{ x_s = (x_s^n)_{n \in \mathbb{N}^*} \in (E^s)^{\mathbb{N}^*} : \|x_s\| = \sup_{n \in \mathbb{N}^*} \kappa^{-n} \|x_s^n\| < +\infty \right\}$$

$$\mathcal{E}^u = \left\{ x_u = (x_u^n)_{n \in \mathbb{N}} \in (E^u)^{\mathbb{N}} : \|x_u\| = \sup_{n \in \mathbb{N}} \kappa^{-n} \|x_u^n\| < +\infty \right\}$$

et munissons les espaces vectoriels produits  $Z = \mathcal{E}^s \times \mathcal{E}^u$  et  $E^s \times Z$  des normes sup.

Considérons l'application  $\Theta : E^s \times Z \rightarrow Z$  définie par  $(x_s^0, x_s, x_u) \mapsto (y_s, y_u)$  où, pour tout  $n \in \mathbb{N}$ , nous posons, en utilisant que  $f = F - T$ ,

$$y_s^{n+1} = F_s(x_s^n + x_u^n) = S(x_s^n) + f_s(x_s^n + x_u^n), \quad (106)$$

$$y_u^n = x_u^n + U^{-1}(x_u^{n+1} - F_u(x_s^n + x_u^n)) = U^{-1}(x_u^{n+1} - f_u(x_s^n + x_u^n)). \quad (107)$$

Notons que si  $(y'_s, y'_u) = \Theta(x_s^0, x'_s, x'_u)$ , nous avons

$$\|y_s^{n+1} - y'_s^{n+1}\| \leq (\|S\| + \delta) \|(x_s^n + x_u^n) - (x_s^n + x'_u^n)\|, \quad (108)$$

donc

$$\|y_s - y'_s\| \leq \kappa^{-1}(\|S\| + \delta) \|(x_s, x_u) - (x'_s, x'_u)\|,$$

et de même

$$\|y_u - y'_u\| \leq \|U^{-1}\|(\kappa + \delta) \|(x_s, x_u) - (x'_s, x'_u)\|.$$

Donc, par la définition de  $\delta'$ , l'application  $\Theta$  est  $\delta'$ -lipschitzienne.

**Lemme 10.27.** *Soit  $k \in (\mathbb{N} - \{0\}) \cup \{+\infty\}$ . Si  $\kappa < 1$  et si  $f$  est de classe  $C^k$ , alors  $\Theta$  est de classe  $C^k$ .*

**Démonstration.** Considérons l'espace de Banach

$$\mathcal{E} = \left\{ x = (x^n)_{n \in \mathbb{N}} \in E^{\mathbb{N}} : \|x\| = \sup_{n \in \mathbb{N}} \kappa^{-n} \|x^n\| < +\infty \right\}.$$

Par la définition de  $\Theta$  dans les formules (106) et (107), et par le fait qu'une application linéaire continue est de classe  $C^\infty$ , il suffit de montrer que l'application  $\tilde{f} : \mathcal{E} \rightarrow \mathcal{E}$  valant  $f$  sur chaque composante (c'est-à-dire définie par  $(x^n)_{n \in \mathbb{N}} \mapsto (f(x^n))_{n \in \mathbb{N}}$ ) est de classe  $C^k$ . Puisque  $\kappa < 1$ , pour tous les éléments  $x = (x^n)_{n \in \mathbb{N}}$  et  $h = (h^n)_{n \in \mathbb{N}}$  de  $\mathcal{E}$ , et pour tout  $t \in [0, 1]$ , les suites  $(x^n)_{n \in \mathbb{N}}$ ,  $(h^n)_{n \in \mathbb{N}}$  et  $(x^n + t h^n)_{n \in \mathbb{N}}$  tendent vers 0 dans  $E$ . Donc les différentielles  $d_{x^n + t h^n}^\ell f$  d'ordre  $\ell$  de  $f$  sont de normes uniformément bornées pour tout  $\ell = 1, \dots, k$ . En particulier, l'application  $\ell$ -linéaire  $D_x^\ell$  de  $\mathcal{E}^\ell$  dans  $\mathcal{E}$  définie par

$$\left( (h_1^n)_{n \in \mathbb{N}}, \dots, (h_\ell^n)_{n \in \mathbb{N}} \right) \mapsto \left( d_{x^n}^\ell f(h_1^n, \dots, h_\ell^n) \right)_{n \in \mathbb{N}}$$

est bien à valeurs dans  $\mathcal{E}$ , et est continue. La formule de Taylor de  $f$  avec reste intégral

$$\begin{aligned} & f(x^n + h^n) \\ &= f(x^n) + \sum_{\ell=1}^{k-1} \frac{1}{\ell!} d_{x^n}^\ell f(h^n, \dots, h^n) + \frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} d_{x^n+th^n}^k f(h^n, \dots, h^n) dt \\ &= f(x^n) + \sum_{\ell=1}^k \frac{1}{\ell!} d_{x^n}^\ell f(h^n, \dots, h^n) \\ &\quad + \frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} (d_{x^n+th^n}^k f - d_{x^n}^k f)(h^n, \dots, h^n) dt, \end{aligned}$$

et le fait que la norme d'opérateur  $\|d_{x^n+th^n}^k f - d_{x^n}^k f\|$  tende vers 0 uniformément en  $n \in \mathbb{N}$  et  $t \in [0, 1]$  quand  $\sup_{n \in \mathbb{N}} \|h^n\|$  tend vers 0, montre de plus que  $\tilde{f}$  est de classe  $C^k$ , ayant pour différentielle d'ordre  $\ell$  l'application  $D_x^\ell$  pour tout  $\ell = 1, \dots, k$ .  $\square$

Par le théorème 10.26, pour tout  $x \in E^s$ , notons  $\varphi(x) \in \mathcal{E}^s \times \mathcal{E}^u$  l'unique point fixe de l'application  $z \mapsto \Theta(x, z)$ , et notons  $G_\kappa : E^s \rightarrow E^u$  l'application qui à  $x \in E^s$  associe  $(\varphi(x))_u^0$  (c'est-à-dire le premier terme de la seconde composante de  $\varphi(x) \in \mathcal{E}^s \times \mathcal{E}^u$ ).

Puisque  $F(0) = T(0) + f(0) = 0$ , le point 0 est un point fixe de l'application  $z \mapsto \Theta(0, z)$  par les formules (106) et (107) définissant  $\Theta$ . Donc par l'unicité dans le théorème 10.26, nous avons  $\varphi(0) = 0$  et  $G_\kappa(0) = 0$ . Par le théorème 10.26, l'application  $\varphi$  est  $\delta'$ -lipschitzienne. Donc  $G_\kappa$  est  $\delta'$ -lipschitzienne (puisque la projection  $(x_s, x_u) \mapsto x_u^0$  de  $\mathcal{E}^s \times \mathcal{E}^u$  dans  $E^u$  est 1-lipschitzienne).

Si  $\kappa < 1$  et s'il existe  $k \geq 1$  tel que  $f$  soit de classe  $C^k$ , alors par le lemme 10.27, l'application  $\Theta$  est de classe  $C^k$ . Donc l'application  $\varphi$  (et par conséquent  $G_\kappa$ ) est de classe  $C^k$  par le théorème 10.26. De plus, si  $d_0 f = 0$ , alors par la formule (105) et le fait que  $(\Theta(x, 0))_u^0 = U^{-1} f_u(x)$  par la formule (107), nous avons  $d_0 G_\kappa = 0$ .

Le but d'avoir introduit l'application  $\Theta$  est que, en regardant les composantes dans la somme directe  $E = E^s \oplus E^u$ , puisque  $x_u^n = y_u^n$  si et seulement si  $x_u^{n+1} = F_u(x_s^n, x_u^n)$  par la formule (107), nous avons l'équivalence

$$\forall n \in \mathbb{N}, F(x_s^n + x_u^n) = x_s^{n+1} + x_u^{n+1} \iff \forall n \in \mathbb{N}, \begin{cases} x_s^{n+1} = y_s^{n+1} \\ x_u^n = y_u^n, \end{cases} \quad (109)$$

c'est-à-dire si et seulement si  $((x_s^n)_{n \in \mathbb{N}^*}, (x_u^n)_{n \in \mathbb{N}})$  est un point fixe dans  $Z$  de l'application  $z \mapsto \Theta(x_s^0, z)$ . Donc pour tout  $x_s \in E^s$ , par la définition de  $\mathcal{E}^u$ , l'élément  $G_\kappa(x_s)$  est un élément  $x_u \in E^u$  tel que

$$x_s + x_u \in W_\kappa^s(F) = \{x \in E : \sup_{n \in \mathbb{N}} \kappa^{-n} \|F^n(x)\| < \infty\}.$$

C'est de plus l'unique tel élément  $x_u$ , par l'unicité du point fixe de  $z \mapsto \Theta(x_s, z)$ . Ceci montre que  $W_\kappa^s(F)$  est le graphe de l'application  $G_\kappa$ .

Soient  $x = x_s + G_\kappa(x_s)$  et  $y = y_s + G_\kappa(y_s)$  deux éléments de  $W_\kappa^s(F)$ . Puisque la norme est adaptée et puisque  $G_\kappa$  est  $\delta'$ -lipschitzienne avec  $\delta' < 1$ , nous avons

$$\|x - y\| = \max\{\|x_s - y_s\|, \|G_\kappa(x_s) - G_\kappa(y_s)\|\} = \|x_s - y_s\|.$$

Donc par l'inclusion immédiate  $F(W_\kappa^s(F)) \subset W_\kappa^s(F)$ , nous avons

$$\|F(x) - F(y)\| = \|F_s(x) - F_s(y)\|.$$

Par les formules (109) et (108), nous avons

$$\|F_s(x) - F_s(y)\| \leq (\|S\| + \delta) \|x - y\|.$$

Donc

$$\forall x, y \in W_\kappa^s(F), \quad \|F(x) - F(y)\| \leq (\|T|_{E^s}\| + \delta) \|x - y\|. \quad (110)$$

Ceci montre la formule (103). L'assertion (104) du théorème 10.25 en découle, car nous avons  $F(0) = 0$  et par l'hypothèse sur  $\delta$ , nous avons  $\|S\| + \delta < \kappa$ .  $\square$

## 10.8.2 Variétés stables et instables locales des difféomorphismes

Soient  $M$  une variété différentielle lisse et  $\pi : TM \rightarrow M$  son fibré tangent. Soient  $\phi : M \rightarrow M$  un  $C^k$ -difféomorphisme (où  $k \in \mathbb{N} - \{0\}$ ) et  $\Lambda$  une partie compacte de  $M$ , invariante par  $\phi$  et  $(\kappa_s, \kappa_u)$ -hyperbolique pour  $\phi$ , où  $0 < \kappa_s < 1 < \kappa_u$ . Reprenons la norme riemannienne  $\|\cdot\|$ , l'application riemannienne  $\exp$  et la distance riemannienne  $d$  de la partie 10.7. Fixons

$$\kappa \in ]\max\{\kappa_s, \kappa_u^{-1}\}, 1[. \quad (111)$$

Munissons le fibré vectoriel  $TM|_\Lambda$  d'une norme  $\|\cdot\|' : TM|_\Lambda \rightarrow \mathbb{R}$  adaptée à  $\phi$ . Une telle norme est utile pour des raisonnements dynamiques, mais le fait qu'elle ne soit en général pas de classe  $C^1$  est la raison de l'utilité d'introduire aussi une norme riemannienne pour la procédure de localisation.

Pour tout  $\epsilon > 0$  assez petit et pour tout  $z \in \Lambda$ , notons

$$V(z, \epsilon) = \exp_z(\{v \in T_z M : \|v\|' < \epsilon\}). \quad (112)$$

Remarquons que puisque deux normes sur  $TM|_\Lambda$  sont équivalentes, et par la formule (98) avec  $x_0 = x = z$ , il existe des constantes  $\epsilon'_0 > 0$  et  $c' \geq 1$  telles que, pour tous les  $\epsilon \in ]0, \epsilon'_0]$  et  $z \in \Lambda$ , nous ayons

$$B_d(z, \frac{1}{c'} \epsilon) \subset V(z, \epsilon) \subset B_d(z, c' \epsilon). \quad (113)$$

Le résultat suivant montre en particulier l'intégrabilité (locale) des champs de plans stable  $(E_x^s)_{x \in \mathbb{R}}$  et instable  $(E_x^u)_{x \in \mathbb{R}}$ , et donne les propriétés dynamiques des sous-variétés intégrales (locales).

**Théorème 10.28.** *Il existe  $c, \epsilon_0 > 0$  tels que pour tout  $x \in \Lambda$ , il existe deux boules ouvertes  $C^k$ -plongées  $W_{\text{loc}}^s(x)$  et  $W_{\text{loc}}^u(x)$ , d'adhérence des boules fermées  $C^k$ -plongées, dans  $M$  contenant  $x$ , appelées respectivement les variétés stables et instables locales de  $x$ , dépendant continûment de  $x$ <sup>230</sup>, telles que*

- (1)  $T_x(W_{\text{loc}}^s(x)) = E_x^s$  et  $T_x(W_{\text{loc}}^u(x)) = E_x^u$ ,
- (2)  $\phi(W_{\text{loc}}^s(x)) \subset W_{\text{loc}}^s(\phi(x))$  et  $\phi^{-1}(W_{\text{loc}}^u(x)) \subset W_{\text{loc}}^u(\phi^{-1}(x))$ ,
- (3)

$$\begin{aligned} \forall n \in \mathbb{N}, \quad \forall y, y' \in W_{\text{loc}}^s(x), \quad d(\phi^n(y), \phi^n(y')) &\leq c \kappa^n d(y, y'), \\ \forall n \in \mathbb{N}, \quad \forall y, y' \in W_{\text{loc}}^u(x), \quad d(\phi^{-n}(y), \phi^{-n}(y')) &\leq c \kappa^n d(y, y'), \end{aligned}$$

230. pour la distance de Hausdorff entre parties d'adhérences compactes de  $M$ , où la distance de Hausdorff entre deux parties  $A$  et  $B$  d'adhérences compactes de  $M$  est la borne inférieure des  $\epsilon > 0$  tels que  $A$  soit contenue dans le  $\epsilon$ -voisinage de  $B$  et que  $B$  soit contenue dans le  $\epsilon$ -voisinage de  $A$

(4)

$$\begin{aligned} W_{\text{loc}}^s(x) &\text{ est égal à } W_{\epsilon_0}^s(x) = \{y \in M : \forall n \in \mathbb{N}, \phi^n(y) \in V(\phi^n(x), \epsilon_0)\}, \\ W_{\text{loc}}^u(x) &\text{ est égal à } W_{\epsilon_0}^u(x) = \{y \in M : \forall n \in \mathbb{N}, \phi^{-n}(y) \in V(\phi^{-n}(x), \epsilon_0)\}. \end{aligned}$$

Nous renvoyons à l'exercice [E.62](#) pour le calcul des variétés stables et instables locales des difféomorphismes d'Anosov linéaire des tores (voir la partie [10.3](#)) et à l'exercice [E.63](#) pour le calcul des variétés stables et instables locales des points de l'attracteur de Smale par l'application du fer à cheval de Smale (voir la partie [10.5](#)).

Remarquons que la notation  $W_{\text{loc}}^s(x)$  et  $W_{\text{loc}}^u(x)$  omet la mention de  $\epsilon$ , ainsi que de choix de la métrique riemannienne  $\|\cdot\|$ , avec son application exponentielle  $\exp$  et sa distance  $d$  associée. Cette notation dépend vraiment de ces choix. Par contre, le germe en  $x$  des sous-variétés  $W_{\text{loc}}^s(x)$  et  $W_{\text{loc}}^u(x)$  ne dépend pas de ces choix, et il est important et suffisant de travailler quitte à restreindre à un voisinage ouvert de  $x$  dans ces sous-variétés.

Remarquons que la dépendance en  $x$  des variétés stables et instables locales est continue, comme affirmé ci-dessus, mais même si  $\Lambda = M$ , elle n'est en général pas très régulière (et en particulier pas forcément de classe  $C^k$  même si  $\Lambda = M$ ). Sauf cadre homogène (comme par exemple pour les feuilles stables et instables du flot géodésique des surfaces hyperboliques, voir la partie [9.5](#)), la seule régularité transverse aux feuilles que nous puissions espérer est la régularité höldérienne (voir par exemple [\[KH\]](#)).

**Démonstration.** Nous montrons l'existence des variétés stables locales, l'existence des variétés instables locales s'en déduit en remplaçant  $\phi$  par  $\phi^{-1}$ .

Notons  $\tilde{F} = \tilde{G}_\phi : TM|_\Lambda \rightarrow TM|_\Lambda$  l'application donnée par le lemme [10.18](#) en prenant  $\psi = \phi$ . En particulier, en rappelant la notation  $\tilde{F}_x = \tilde{F}|_{T_x M}$  pour tout  $x \in \Lambda$ , nous avons

$$\tilde{F}_x(0) = \exp_{\phi(x)}^{-1} \circ \phi \circ \exp_x(0) = \exp_{\phi(x)}^{-1}(\phi(x)) = 0. \quad (114)$$

Si la constante  $\theta$  apparaissant dans ce lemme est assez petite, alors pour tout  $x \in \Lambda$ , l'application  $\tilde{F}_x - T_x \phi$  est  $\delta$ -lipschitzienne avec  $\delta > 0$  indépendant de  $x$  et que nous pouvons supposer arbitrairement petit.

Le résultat technique suivant est juste une version fibrée du théorème [10.25](#). Rappelons que  $\kappa < 1$  a été fixé dans la formule [\(111\)](#).

**Lemme 10.29.** *Il existe une famille continue  $(g_x)_{x \in \Lambda}$  d'applications  $g_x : E_x^s \rightarrow E_x^u$  de classe  $C^k$  et  $\delta'$ -lipschitziennes avec  $\delta' < 1$ , telle que  $g_x(0) = 0$  et  $d_0 g_x = 0$ , et dont le graphe est*

$$\widetilde{W}_\kappa^s(x, \tilde{F}) = \{v \in T_x M : \sup_{n \in \mathbb{N}} \kappa^{-n} \|\tilde{F}^n(v)\|' < +\infty\}. \quad (115)$$

De plus, nous avons

$$\tilde{F}_x(\widetilde{W}_\kappa^s(x, \tilde{F})) \subset \widetilde{W}_\kappa^s(\phi(x), \tilde{F}), \quad (116)$$

$$\forall v, v' \in \widetilde{W}_\kappa^s(x, \tilde{F}), \quad \|\tilde{F}(v) - \tilde{F}(v')\|' \leq (\kappa_s + \delta) \|v - v'\|', \quad (117)$$

$$\forall v \in \widetilde{W}_\kappa^s(x, \tilde{F}), \quad \lim_{n \rightarrow +\infty} \kappa^{-n} \|\tilde{F}^n(v)\|' = 0. \quad (118)$$

**Démonstration.** Notons  $T = \phi_*$  l'application définie par la formule [\(94\)](#) et

$$F : \sigma \mapsto \{x \mapsto \tilde{F}_{\phi^{-1}(x)}(\sigma(\phi^{-1}(x)))\}, \quad (119)$$

qui sont des endomorphismes linéaires continus de l'espace de Banach  $\Gamma_b(TM|_\Lambda)$  des sections ensemblistes bornées (pas forcément continues) du fibré vectoriel  $TM|_\Lambda$ . Par l'équivalence des définitions 10.11, l'opérateur  $T$  est hyperbolique, et a pour décomposition associée  $\Gamma_b(TM|_\Lambda) = \Gamma_b(E^s) \oplus \Gamma_b(E^u)$ , où  $E^s = \bigcup_{x \in \Lambda} E_x^s$  et  $E^u = \bigcup_{x \in \Lambda} E_x^u$  sont les sous-fibrés stables et instables de  $TM|_\Lambda$ . Le fait que la norme  $\|\cdot\|'$  de  $TM|_\Lambda$  soit adaptée pour  $\phi$  montre que la norme uniforme  $\|\sigma\|'_\infty = \sup_{x \in \Lambda} \|\sigma(x)\|'$  sur  $\Gamma_b(TM|_\Lambda)$  est adaptée pour  $T$ , et que la constante d'hyperbolicité  $\text{ch}(T)$  de  $T$  est au plus  $\max\{\kappa_s, \kappa_u^{-1}\}$ . En particulier, nous avons  $\kappa \in ]\text{ch}(T), 1[$  par la formule (111). Nous avons  $F(0) = 0$  par les formules (119) et (114). Par les explications du début de démonstration du théorème 10.28, l'application  $F - T$  est  $\delta$ -lipschitzienne avec  $\delta > 0$  que nous pouvons supposer strictement inférieur à  $\min\{\kappa - \text{ch}(T), \text{ch}(T)^{-1} - \kappa\} > 0$ .

Puisque  $\phi$  est de classe  $C^k$ , par les propriétés de la procédure de localisation (lemme 10.18), l'application  $\tilde{F} = \tilde{G}_\phi$  est la restriction à  $TM|_\Lambda$  d'une application de classe  $C^k$  sur  $TM$ . Par la dernière affirmation du lemme 10.18, nous avons  $d_0F = T$ , donc  $d_0(F - T) = 0$  par linéarité de  $T$ .

Par le théorème 10.25, il existe donc une application  $G_\kappa : \Gamma_b(E^s) \rightarrow \Gamma_b(E^u)$ , qui est  $\delta'$ -lipschitzienne avec  $\delta' < 1$  et de classe  $C^k$ , telle que  $G_\kappa(0) = 0$ ,  $d_0G_\kappa = 0$  (car  $d_0(F - T) = 0$ ) et par la dernière affirmation du théorème 10.25) et  $W_\kappa^s(F)$  soit le graphe de  $G_\kappa$ .

Pour tous les  $x \in \Lambda$  et  $v \in T_xM$ , considérons la section ensembliste bornée  $\sigma_v$  de  $TM|_\Lambda$  valant 0 en tout point de  $\Lambda$  différent de  $x$ , et valant  $v$  en  $x$ . Pour tout  $x \in M$  fixé, l'application de  $T_xM$  dans  $\Gamma_0(TM|_\Lambda)$  définie par  $v \mapsto \sigma_v$  est linéaire continue, et en particulier, nous avons

$$d_0((\sigma_v)|_{T_xM}) = (\sigma_v)|_{T_xM} .$$

Nous avons de plus  $\|\sigma_v\|'_\infty = \|v\|'$  et  $F^n(\sigma_v) = \sigma_{\tilde{F}^n(v)}$  pour tout  $n \in \mathbb{N}$ . En particulier,

$$v \in \tilde{W}_\kappa^s(\pi(v), \tilde{F}) \iff \sigma_v \in W_\kappa^s(F) . \quad (120)$$

Pour tout  $x \in \Lambda$ , notons  $g_x : E_x^s \rightarrow E_x^u$  l'application définie par

$$g_x : v \mapsto G_\kappa(\sigma_v)(x) ,$$

en remarquant que  $\sigma_v$  appartient à  $\Gamma_b(E^s)$  et  $G_\kappa(\sigma_v)$  à  $\Gamma_b(E^u)$  lorsque  $v \in E_x^s$ .

Pour tout  $x \in \Lambda$ , l'application  $v \mapsto \sigma_v$  de  $T_xM$  dans  $\Gamma_b(TM|_\Lambda)$ , ainsi que l'application de  $\Gamma_b(TM|_\Lambda)$  dans  $T_xM$  d'évaluation des sections en  $x$ , sont linéaires continues, donc de classe  $C^k$ . Puisque l'application  $G_\kappa$  est de classe  $C^k$  et par le théorème de dérivation des fonctions composées, l'application  $g_x : E_x^s \rightarrow E_x^u$  est donc de classe  $C^k$ . Puisque  $d_0G_\kappa = 0$ , nous avons  $d_0g_x = 0$  pour tout  $x \in \Lambda$ .

Montrons que le graphe de  $g_x$  est égal à  $\tilde{W}_\kappa^s(x, \tilde{F})$ . Pour tous les  $x \in \Lambda$  et  $\sigma \in \Gamma_b(TM|_\Lambda)$ , nous avons  $\tilde{F}(\sigma(x)) = F(\sigma)(\phi(x))$  par la définition 119 de  $F$ , et donc par récurrence

$$\tilde{F}^n(\sigma(x)) = F^n(\sigma)(\phi^n(x))$$

pour tout  $n \in \mathbb{N}$ . Par conséquent, si  $\sigma \in W_\kappa^s(F)$ , alors

$$\sup_{n \in \mathbb{N}} \kappa^{-n} \|\tilde{F}^n(\sigma(x))\|' = \sup_{n \in \mathbb{N}} \kappa^{-n} \|F^n(\sigma)(\phi^n(x))\|' \leq \sup_{n \in \mathbb{N}} \kappa^{-n} \|F^n(\sigma)\|'_\infty < +\infty ,$$

donc  $\sigma(x) \in \tilde{W}_\kappa^s(x, \tilde{F})$ . Pour tout  $v \in E_x^s$ , nous avons

$$v + g_x(v) = (\sigma_v + G_\kappa(\sigma_v))(x)$$

et  $\sigma_v + G_\kappa(\sigma_v) \in W_\kappa^s(F)$  puisque  $W_\kappa^s(F)$  est le graphe de l'application  $G_\kappa$ . Nous avons donc

$$v + g_x(v) \in \widetilde{W}_\kappa^s(x, \widetilde{F}) .$$

Réciproquement, si un vecteur  $w = w_s + w_u \in T_x M$  appartient à  $\widetilde{W}_\kappa^s(x, \widetilde{F})$ , alors nous avons  $\sigma_{w_s} \in \Gamma_b(E^s)$ ,  $\sigma_{w_u} \in \Gamma_b(E^u)$  et la section  $\sigma_w = \sigma_{w_s} + \sigma_{w_u}$  appartient à  $W_\kappa^s(F)$  par la formule (120). Puisque  $W_\kappa^s(F)$  est le graphe de  $G_\kappa$ , nous avons donc  $\sigma_{w_u} = G_\kappa(\sigma_{w_s})$ , et

$$w_u = \sigma_{w_u}(x) = G_\kappa(\sigma_{w_s})(x) = g_x(w_s) .$$

Donc  $w$  appartient au graphe de  $g_x$ .

Le fait que  $g_x$  soit  $\delta'$ -lipschitzienne découle du fait que  $G_\kappa$  l'est. Montrons que  $g_x$  dépend continûment de  $x$ . Il suffit de montrer que si  $\sigma_s \in \Gamma_b^0(E^s)$ , alors l'application  $\sigma_u \in \Gamma_b(E^s)$  définie par  $x \mapsto g_x(\sigma_s(x))$  est continue. Notons  $T^0$  et  $F^0$  les restrictions de  $T$  et  $F$  au sous-espace de Banach  $\Gamma_b^0(TM|_\Lambda)$ . Par le théorème 10.25 appliqué à  $T^0$  et  $F^0$ , il existe une application  $G_\kappa^0$  dont le graphe est  $W_\kappa^s(F^0)$ . Ceci implique comme ci-dessus que  $(\sigma_s + G_\kappa^0(\sigma_s))(x)$  appartient à  $\widetilde{W}_\kappa^s(x, \widetilde{F})$ , donc que

$$\sigma_u(x) = g_x(\sigma_s(x)) = G_\kappa^0(\sigma_s)(x) .$$

Par conséquent,  $\sigma_u = G_\kappa^0(\sigma_s)$  est continue.

La formule (117) découle du théorème 10.25 (et plus précisément de la formule (103)).

L'anté-pénultième assertion (116) est immédiate par définition des  $\widetilde{W}_\kappa^s(x, \widetilde{F})$ .

En itérant la formule (117) avec  $v' = 0$ , la dernière assertion (118) découle du fait que  $\kappa > \kappa_s + \delta$  et que pour tout  $x \in \Lambda$ , nous avons  $\widetilde{F}(v') = 0$  si  $v' = 0 \in T_x M$ . Ceci conclut la démonstration du lemme technique 10.29.  $\square$

Revenons à la démonstration du théorème 10.28. Fixons  $\epsilon_0 > 0$  assez petit (à préciser ultérieurement), et pour tous les  $\epsilon \in ]0, \epsilon_0]$  et  $x \in \Lambda$ , posons

$$W_\epsilon^s(x) = \{y \in M : \forall n \in \mathbb{N}, \phi^n(y) \in V(\phi^n(x), \epsilon)\} , \quad (121)$$

$B_x^s(\epsilon) = \{v \in E_x^s : \|v\|' < \epsilon\}$  et  $j_x : B_x^s(\epsilon_0) \rightarrow M$  l'application définie par

$$j_x : v \mapsto \exp_x(v + g_x(v)) ,$$

où  $(g_x)_{x \in X}$  est la famille donnée par le lemme 10.29. L'application  $j_x$  dépend continûment de  $x$ . Le second lemme technique suivant décrit les propriétés de cette application  $j_x$ .

**Lemme 10.30.** (i) Si  $\epsilon_0$  est assez petit, l'application  $j_x$  est un plongement  $C^k$ , envoyant 0 sur  $x$ , dont l'image est  $W_{\epsilon_0}^s(x)$ , et qui vérifie  $T_0 j_x(E_x^s) = E_x^s$ .

(ii) Nous avons  $\phi(W_\epsilon^s(x)) \subset W_{\kappa\epsilon}^s(\phi(x))$ .

**Démonstration.** (i) L'application de  $E_x^s$  dans  $E_x^s \oplus E_x^u$ , définie par  $v_s \mapsto v_s + g_x(v_s)$ , est un  $C^k$ -plongement comme graphe d'une application de classe  $C^k$ , qui envoie 0 sur  $(0, 0)$ . Puisque  $g_x$  est  $\delta'$ -lipschitzienne, si  $\epsilon_0$  est assez petit, l'image de  $B_x^s(\epsilon_0)$  par cette application est contenue dans le domaine de définition du difféomorphisme  $\exp_x$ . Donc  $j_x$  est un  $C^k$ -plongement qui envoie 0 sur  $\exp_x 0 = x$ , si  $\epsilon_0$  est assez petit.

L'application tangente

$$T_0 j_x : T_0(B_x^s(\epsilon)) = E_x^s \rightarrow T_{j_x(0)} M = T_x M$$

de  $j_x$  en 0 est l'application  $v_s \mapsto v_s$  car  $d_0 g_x = 0$  par le lemme 10.29 et  $T_0 \exp_x = \text{id}_{T_x M}$ . Nous avons donc  $T_0 j_x(E_x^s) = E_x^s$ .

Puisque  $\tilde{F}_x$  coïncide avec  $\exp_{\phi(x)}^{-1} \circ \phi \circ \exp_x$  sur des petits vecteurs (uniformément en  $x \in \Lambda$  d'après le lemme 10.18) et par la définition des voisinages  $V(z, \epsilon)$  dans la formule 112, si  $\epsilon_0$  est assez petit, pour tout  $\epsilon \in ]0, \epsilon_0]$ , nous avons

$$\begin{aligned} \exp_x^{-1}(W_\epsilon^s(x)) &= \bigcap_{n \in \mathbb{N}} \exp_x^{-1}(\phi^{-n} \circ \exp_{\phi^n(x)}(\{v \in T_{\phi^n(x)} M : \|v\|' < \epsilon\})) \\ &= \bigcap_{n \in \mathbb{N}} \tilde{F}^{-n}(\{v \in T_{\phi^n(x)} M : \|v\|' < \epsilon\}). \end{aligned}$$

Cette intersection infinie est contenue dans l'intersection  $\tilde{W}_\kappa^s(x, \tilde{F}) \cap \{v \in T_x M : \|v\|' < \epsilon\}$  par la définition de  $\tilde{W}_\kappa^s(x, \tilde{F})$  dans la formule (115) et par le fait que  $\tilde{F}$  est  $\kappa$ -lipschitzien sur cette intersection infinie par l'assertion (103) du théorème 10.25 appliquée avec  $\kappa = 1$ . Puisque  $\tilde{F}(0) = 0$  et  $\tilde{F}$  contracte  $\tilde{W}_\kappa^s(x, \tilde{F})$ , l'inclusion inverse est aussi vérifiée, et donc

$$W_\epsilon^s(x) = \exp_x(\tilde{W}_\kappa^s(x, \tilde{F}) \cap \{v \in T_x M : \|v\|' < \epsilon\}). \quad (122)$$

Puisque le graphe de  $g_x$  est égal à  $\tilde{W}_\kappa^s(x, \tilde{F})$  par le lemme 10.29 et puisque la norme est adaptée, nous avons  $j_x(B_x^s(\epsilon_0)) = W_{\epsilon_0}^s(x)$ .

(ii) Puisque  $\kappa \geq \kappa_s + \delta$ , le fait que  $\phi(W_\epsilon^s(x)) \subset W_{\kappa\epsilon}^s(\phi(x))$  découle du fait que, par le lemme 10.29, l'application  $\tilde{F}$  est  $(\kappa_s + \delta)$ -lipschitzienne sur  $\tilde{W}_\kappa^s(x, \tilde{F})$ .  $\square$

Vérifions maintenant les assertions du théorème 10.28 concernant les variétés stables locales. L'assertion (4) est la définition de  $W_{\text{loc}}^s(x) = W_{\epsilon_0}^s(x)$ .

Le lemme 10.30 (i) montre que  $T_0 j_x(E_x^s) = E_x^s$  et que  $W_{\text{loc}}^s(x)$  est l'image de la boule ouverte  $B_x^s(\epsilon_0)$  de  $E_x^s$  par le plongement  $j_x$ . L'assertion  $T_x W_{\text{loc}}^s(x) = E_x^s$  du théorème 10.28 (1) en découle.

L'assertion  $\phi(W_{\text{loc}}^s(x)) \subset W_{\text{loc}}^s(\phi(x))$  du théorème 10.28 (2) découle du lemme 10.30 (ii).

Enfin, l'assertion (3) concernant les variétés stables locales découle de la propriété (117) de contraction de  $\tilde{F}$ , qui coïncide avec  $\exp_{\phi(x)}^{-1} \circ \phi \circ \exp_x$  sur des vecteurs proches de la section nulle, de la formule (98) et de l'équivalence des normes.  $\square$

### 10.8.3 Variétés stables et instables globales des difféomorphismes

Reprenons les notations  $M$ ,  $\phi$ ,  $\Lambda$  et  $d$  pour la définition des variétés stables et instables locales.

Pour tout point  $x \in \Lambda$ , définissons la *variété stable* (globale) de  $x$ , notée  $W^s(x)$  ou  $W^s(x, \phi)$  lorsqu'il convient de préciser  $\phi$ , par

$$W^s(x) = \{y \in M : \lim_{n \rightarrow +\infty} d(\phi^n(x), \phi^n(y)) = 0\}. \quad (123)$$

De même, la *variété instable* (globale) de  $x$ , notée  $W^u(x)$  ou  $W^u(x, \phi)$  lorsqu'il convient de préciser  $\phi$ , est définie par

$$W^u(x) = W^s(x, \phi^{-1}) = \{y \in M : \lim_{n \rightarrow +\infty} d(\phi^{-n}(x), \phi^{-n}(y)) = 0\}. \quad (124)$$

Les variétés stables et instables des points de  $\Lambda$  sont clairement équivariantes par  $\phi$  :

$$\phi(W^s(x)) = W^s(\phi(x)) \quad \text{et} \quad \phi(W^u(x)) = W^u(\phi(x)) . \quad (125)$$

Notons que deux variétés stables sont ou bien disjointes ou bien confondues : pour tous les  $x, y \in \Lambda$ , si  $W^s(x) \cap W^s(y) \neq \emptyset$ , alors  $W^s(x) = W^s(y)$ .<sup>231</sup> De même, deux variétés instables sont ou bien disjointes ou bien confondues.

Par la définition des variétés stables et instables, et par le théorème 10.28, nous avons<sup>232</sup>

$$W^s(x) = \bigcup_{n \in \mathbb{N}} \phi^{-n}(W_{\text{loc}}^s(\phi^n(x))) \quad \text{et} \quad W^u(x) = \bigcup_{n \in \mathbb{N}} \phi^n(W_{\text{loc}}^u(\phi^{-n}(x))) . \quad (126)$$

Par le théorème 10.28 (2), ces unions sont des unions croissantes. Toujours par le théorème 10.28, il existe  $\epsilon_2 > 0$  tel que pour tout  $x \in \Lambda$ , nous ayons

$$W^s(x) = \left\{ y \in M : \limsup_{n \rightarrow +\infty} d(\phi^n(x), \phi^n(y)) \leq \epsilon_2 \right\} ,$$

et il existe  $\kappa > 0$  tel que pour tout  $x \in \Lambda$ , nous ayons

$$W^s(x) = \left\{ y \in M : \lim_{n \rightarrow +\infty} \kappa^{-n} d(\phi^n(x), \phi^n(y)) = 0 \right\} ,$$

et de même pour les variétés instables en remplaçant  $\phi$  par  $\phi^{-1}$ .

Le résultat suivant dit en particulier que les variétés stables et instables globales sont des sous-variétés immergées (qui ne sont pas fermées en général, voir par exemple l'exercice E.62).

**Théorème 10.31.** *Pour tout point  $x \in \Lambda$ , il existe des immersions de classe  $C^k$  injectives  $\bar{j}_x^s : E_x^s \rightarrow M$  et  $\bar{j}_x^u : E_x^u \rightarrow M$  d'images égales à  $W^s(x)$  et  $W^u(x)$  respectivement, telles que les applications  $x \mapsto \bar{j}_x^s$  et  $x \mapsto \bar{j}_x^u$  soient continues sur  $\Lambda$  pour la topologie  $C^k$ . Le sous-espace tangent en  $x$  à sa variété stable/instable est le sous-espace stable/instable de l'espace tangent en  $x$  :*

$$T_x(W^s(x)) = E_x^s \quad \text{et} \quad T_x(W^u(x)) = E_x^u .$$

Nous renvoyons à l'exercice E.62 pour le calcul des variétés stables et instables globales des difféomorphismes d'Anosov linéaire des tores (voir la partie 10.3) et à l'exercice E.63 pour le calcul des variétés stables et instables globales des points de l'attracteur de Smale par l'application du fer à cheval de Smale (voir la partie 10.5).

**Démonstration.** Il suffit de montrer le résultat pour la variété stable. Reprenons les notations  $\tilde{F}$  et  $(g_x)_{x \in \Lambda}$  de la démonstration du théorème 10.28. Pour tous les  $x \in \Lambda$  et  $v_0 \in E_x^s$ , soit  $V$  un voisinage compact de  $v_0$  dans  $E_x^s$ . Puisque  $\tilde{F}_x$  coïncide avec  $\exp_{\phi(x)}^{-1} \circ \phi \circ \exp_x$

231. En effet, si  $z \in W^s(x) \cap W^s(y)$ , alors pour tout  $w \in W^s(x)$ , la distance  $d(\phi^n(w), \phi^n(y))$ , qui est inférieure ou égale à la somme  $d(\phi^n(w), \phi^n(x)) + d(\phi^n(x), \phi^n(z)) + d(\phi^n(z), \phi^n(y))$  par l'inégalité triangulaire, converge vers 0 quand  $n \rightarrow +\infty$ , donc  $w \in W^s(y)$ . D'où l'inclusion  $W^s(x) \subset W^s(y)$ , qui est une égalité par symétrie.

232. L'assertion (4) du théorème 10.28 et la formule (113) montrent que si  $y \in W^s(x)$ , alors il existe  $n_0 \in \mathbb{N}$  tel que pour tout  $n \geq n_0$ , nous avons  $\phi^n(y) \in V(\phi^n(x), \epsilon_0)$ . Donc  $\phi^{n_0}(y) \in W_{\text{loc}}^s(\phi^{n_0}(x))$ , ce qui donne l'inclusion  $W^s(x) \subset \bigcup_{n \in \mathbb{N}} \phi^{-n}(W_{\text{loc}}^s(\phi^n(x)))$ . L'assertion (3) du théorème 10.28 montre l'inclusion réciproque.



sur les petits vecteurs de  $T_x M$  par le lemme 10.18, puisque le graphe de  $g_x$  est égal à  $\widetilde{W}_\kappa^s(x, \widetilde{F})$  (défini par la formule (115)) sur lequel  $\widetilde{F}$  contracte par le lemme 10.29, il existe  $n_0 \in \mathbb{N}$  tel que pour tous les  $n \geq n_0$  et  $v \in V$ , le point

$$\bar{j}_x^s(v) = \phi^{-n} \circ \exp_{\phi^n(x)} \widetilde{F}^n(v + g_x(v))$$

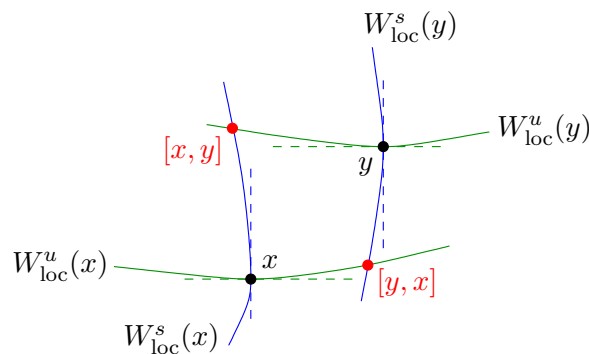
ne dépend pas de  $n$ . En utilisant la partie gauche de la formule (126) et les propriétés des variétés stables locales, le résultat en découle.  $\square$

## 10.9 La structure de produit local

Reprenons les notations  $M$ ,  $\phi$ ,  $\Lambda$  et  $d$  pour la définition des variétés stables et instables (locales et globales).

Par la transversalité des sous-fibrés stable et instable  $E^s$  et  $E^u$ , et par la continuité des feuilles stables et instables locales, qui sont tangentes aux fibres stables et instables, si  $x$  et  $y$  sont suffisamment proches, alors la feuille stable locale  $W_{\text{loc}}^s(x)$  de  $x$  rencontre la feuille instable locale  $W_{\text{loc}}^u(y)$  de  $y$  en un et un seul point, noté  $[x, y]$  (voir le dessin ci-dessous) et appelé le *crochet de Bowen* (local) de  $x$  et  $y$ . Plus précisément, par uniformité, il existe  $\epsilon_0'' = \epsilon_0''(\Lambda) > 0$  et  $\theta_0 = \theta_0(\Lambda) > 1$  tels que pour tous les  $x, y \in \Lambda$  et  $\epsilon \in ]0, \epsilon_0]$ , si  $d(x, y) \leq \theta_0^{-1}\epsilon$ , alors

- l'intersection  $W_{\epsilon_0}^s(x) \cap W_{\epsilon_0}^u(y)$  est réduite à un seul point, et
- ce point appartient de plus à  $W_\epsilon^s(x) \cap W_\epsilon^u(y)$ .



Il découle de la définition (121) des variétés stables locales et de la propriété (113) de l'exponentielle des boules pour la norme adaptée, qu'il existe une constante  $c_1 > 0$  telle que pour tous les  $\epsilon > 0$  assez petit et  $x, y \in \Lambda$  tels que  $d(x, y) \leq \epsilon$ , nous ayons

$$[x, y] \in B_d(x, c_1\epsilon) \cap B_d(y, c_1\epsilon). \quad (127)$$

Par invariance et unicité, si  $d(x, y) \leq \theta_0^{-1}\epsilon_0$  et  $d(\phi(x), \phi(y)) \leq \theta_0^{-1}\epsilon_0$ , alors

$$\phi([x, y]) = [\phi(x), \phi(y)].$$

Nous dirons que la partie hyperbolique  $\Lambda$  admet une structure de *produit local* s'il existe  $\epsilon_1 \in ]0, \epsilon_0[$  tel que pour tous les  $x, y \in \Lambda$  tels que  $d(x, y) \leq \theta_0^{-1}\epsilon_1$ , nous ayons  $[x, y] \in \Lambda$ . Ceci est automatiquement vérifié si  $\Lambda = M$ , c'est-à-dire si  $\Phi$  est un difféomorphisme d'Anosov. Nous montrerons dans l'exercice E.63 que cette propriété est vérifiée pour l'attracteur de Smale  $\Lambda$  (voir la partie 10.5). Mais cette propriété n'est pas toujours vérifiée.

Si  $\Lambda$  admet une structure de produit local, alors pour tous les  $\epsilon > 0$  assez petit et  $x \in \Lambda$ , nous avons un homéomorphisme de  $\Lambda \cap B_d(x, \epsilon)$  dans  $V_x^s \times V_x^u$  où  $V_x^s$  est un voisinage de  $x$  dans  $\Lambda \cap W_{\text{loc}}^s(x)$  et  $V_x^u$  est un voisinage de  $x$  dans  $\Lambda \cap W_{\text{loc}}^u(x)$ , donné par  $y \mapsto ([x, y], [y, x])$ .

Nous dirons que la partie hyperbolique  $\Lambda$  est *localement maximale* (ou isolée) s'il existe un voisinage  $U$  de  $\Lambda$  tel que  $\Lambda = \bigcap_{n \in \mathbb{Z}} \phi^{-n}(U)$ .

**Théorème 10.32.** *Les assertions suivantes sont équivalentes :*

- (1) la partie hyperbolique  $\Lambda$  admet une structure de produit local,
- (2) la partie hyperbolique  $\Lambda$  est localement maximale,
- (3) il existe  $\epsilon_2 > 0$  et un voisinage  $U$  de  $\Lambda$  tel que pour toute  $\epsilon_2$ -pseudo-orbite  $(x_i)_{i \in \mathbb{Z}}$  de  $\phi$  dans  $U$  et pour tout  $x \in M$  tel que  $\sup_{i \in \mathbb{Z}} d(\phi^i(x), x_i) \leq c_0 \epsilon_2$  (où  $c_0 > 0$  est la constante du lemme de pistage 10.20 pour  $\Lambda$ ), nous ayons  $x \in \Lambda$ .

**Démonstration.** Montrons que (3) implique (2). Prenons un voisinage  $U$  de  $\Lambda$  assez petit pour vérifier les hypothèses du lemme de pistage et de l'assertion (3). Alors pour tout  $x \in \bigcap_{n \in \mathbb{Z}} \phi^{-n}(U)$ , l'orbite  $(\phi^n(x))_{n \in \mathbb{Z}}$  est une  $\epsilon$ -pseudo-orbite dans  $U$  qui reste à distance nulle (donc au plus  $c_0 \epsilon$ ) de l'orbite de  $x$  pour tout  $\epsilon > 0$ . Donc l'unicité dans le lemme de pistage et l'assertion (3) impliquent que  $x \in \Lambda$ , donc  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U) \subset \Lambda$ , et l'inclusion inverse est évidente par invariance.

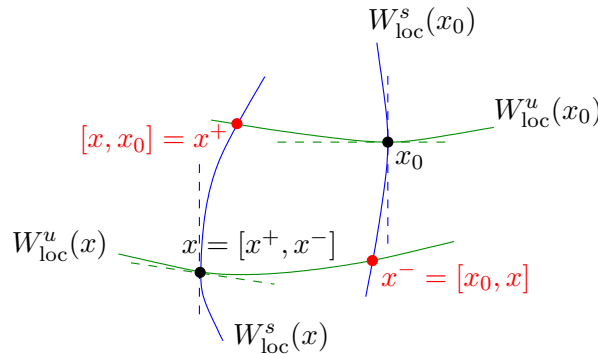
Montrons que (2) implique (1). Soit  $U$  un voisinage de  $\Lambda$  tel que  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U) = \Lambda$ . Soit  $\epsilon \in ]0, \epsilon_0]$  tel que les feuilles locales  $W_\epsilon^s(x)$  et  $W_\epsilon^u(x)$  soient contenues dans  $U$  pour tout  $x \in \Lambda$ .

Soient  $x, y \in \Lambda$  tels que  $d(x, y) \leq \theta_0^{-1} \epsilon$ . Alors par la définition du crochet de Bowen, nous avons  $[x, y] \in W_\epsilon^s(x) \cap W_\epsilon^u(y) \subset U$ . Pour tout  $n \in \mathbb{N}$ , par les propriétés du théorème 10.28 (2) des feuilles locales et par l'invariance de  $\Lambda$  par  $\phi$ , nous avons

$$\phi^n([x, y]) \in W_\epsilon^s(\phi^n(x)) \subset U \quad \text{et} \quad \phi^{-n}([x, y]) \in W_\epsilon^u(\phi^{-n}(x)) \subset U.$$

D'où  $[x, y] \in \bigcap_{n \in \mathbb{Z}} \phi^{-n}(U) = \Lambda$ .

Montrons que (1) implique (3). Soient  $\epsilon_1 > 0$  comme dans la définition de la propriété de produit local. Soient  $(x_i)_{i \in \mathbb{Z}}$  et  $x$  comme dans (3), pour des  $\epsilon_2 > 0$  et  $U$  assez petit que nous allons préciser. Quitte à prendre  $U$  assez petit et à remplacer  $x_i$  par un de ses points le plus proche sur  $\Lambda$ , nous pouvons supposer que  $x_i \in \Lambda$ . Par la dernière affirmation du théorème 10.20, si  $\epsilon_2 > 0$  et  $U$  sont assez petit, nous pouvons supposer que  $x$  appartient à un ensemble hyperbolique compact  $\Lambda'$  contenant  $\Lambda$ .



Si  $\epsilon_2 > 0$  est assez petit (devant les constantes nécessaires pour définir le crochet de Bowen pour la partie hyperbolique compacte  $\Lambda'$ ), alors par l'hypothèse de l'assertion (3), nous avons

$$d(x, x_0) \leq c_0 \epsilon_0 \leq \theta_0(\Lambda')^{-1} \epsilon_0''(\Lambda').$$

Donc les points  $x^+ = [x, x_0]$  et  $x^- = [x_0, x]$  sont bien définis. De plus, par la formule (127), nous avons  $x^+, x_- \in B_d(x_0, c_1 c_0 \epsilon_2)$ . Par l'inégalité triangulaire, si  $\epsilon_2 > 0$  est assez petit, nous avons donc

$$d(x_-, x_+) \leq 2c_1 c_0 \epsilon_2 \leq \theta_0(\Lambda)^{-1} \epsilon_1.$$

Montrons qu'ils appartiennent à  $\Lambda$ , ce qui par l'assertion (1) et la formule centrée ci-dessus implique que  $x$ , qui est égal à  $[x^+, x^-]$ , appartient à  $\Lambda$ . Par symétrie, il suffit de montrer que  $x^+ \in \Lambda$ .

Montrons par récurrence que si  $\epsilon_2 > 0$  est assez petit, il existe une constante  $c_2 > 0$  et une suite  $(x'_i)_{i \in \mathbb{N}}$  telle que  $x'_0 = x_0$ ,  $x'_i \in \Lambda \cap W_{c_2 \epsilon_2}^s(x_i)$  et  $x'_{i+1} = [x_{i+1}, \phi(x'_i)]$ .

Nous avons  $x_0 \in \Lambda \cap W_{c_2 \epsilon_2}^s(x_0)$  pour tout  $c > 0$  et nous posons  $x'_0 = x_0$ . Supposons  $x'_i$  construit. Alors pour  $\epsilon_2$  est assez petit devant la constante  $\epsilon_0$  du théorème 10.28 (plus précisément si  $\epsilon_2 \leq \frac{\epsilon_0}{c_2}$ ), nous avons  $\phi(x'_i) \in V(\phi(x_i), c_2 \epsilon_2)$  par la définition (121) des variétés stables locales. Si  $\epsilon_2$  est assez petit devant la constante  $\epsilon'_0$  de la propriété (113) de l'exponentielle des boules pour la norme adaptée (plus précisément si  $\epsilon_2 \leq \frac{\epsilon'_0}{c_2}$ ), alors cette propriété montre que

$$d(\phi(x_i), \phi(x'_i)) \leq c'(c_2 \epsilon_2).$$

Donc, si  $\epsilon_2$  est assez petit, puisque la suite  $(x_i)_{i \in \mathbb{Z}}$  est une  $\epsilon_2$ -pseudo-orbite de  $\phi$ , nous avons

$$d(x_{i+1}, \phi(x'_i)) \leq d(x_{i+1}, \phi(x_i)) + d(\phi(x_i), \phi(x'_i)) \leq \epsilon_2 + c'(c_2 \epsilon_2) \leq \theta_0^{-1} \epsilon_1.$$

Par l'assertion (1), le point  $x'_{i+1} = [x_{i+1}, \phi(x'_i)]$  est donc bien défini et appartient à  $\Lambda$ . Comme affirmé dans [Yoc2, page 284], il appartient à  $W_{c_2 \epsilon_2}^s(x_{i+1})$  si  $c_2$  est assez grand (indépendant de  $i$ ).

Notons alors  $(x'_i)_{i \in \mathbb{Z}}$  la suite dans  $\Lambda$  étendue par  $x'_i = x_i$  pour  $i \neq 0$ . Si  $\epsilon_2$  est assez petit et  $i \geq 0$ , nous avons  $\phi^{-1}(W_{\epsilon'_0}^u(x'_{i+1})) \subset W_{\epsilon'_0}^u(x'_i)$ . Donc  $\phi^i(x^+) = [\phi^i(x), x'_i] \in \Lambda$ . Par conséquent  $x^+ = \lim_{i \rightarrow +\infty} \phi^{-i}(x'_i)$ , qui appartient à  $\Lambda$ , comme voulu.  $\square$

Nous dirons qu'une partie  $\Lambda$  d'une variété  $M$  munie d'un  $C^1$ -difféomorphisme  $\phi$  est une *pièce basique* de  $\phi$  si elle est invariante, compacte, hyperbolique, localement maximale et si  $\phi$  est transitif sur  $\Lambda$  (c'est-à-dire admet une orbite dense dans  $\Lambda$ ).

Par exemple, si  $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  est le fer à cheval de Smale, alors l'attracteur de Smale  $\Lambda$  est une pièce basique : par construction, le carré unité  $Q = ]0, 1[^2$  est un voisinage de  $\Lambda$  tel que  $\Lambda = \bigcap_{n \in \mathbb{Z}} \phi^n(Q)$ , donc  $\Lambda$  est localement maximal. Voir l'exercice E.63 pour une vérification directe que  $\Lambda$  vérifie la propriété de produit local.

Le fait que les points de l'attracteur de Smale, qui sont périodiques pour le fer à cheval, sont denses dans cet attracteur est en fait un résultat général sur les pièces basiques.

**Proposition 10.33.** *Si  $\Lambda$  est une pièce basique d'un  $C^1$ -difféomorphisme  $\phi$  d'une variété  $M$ , alors l'ensemble des points périodiques de  $\phi$  dans  $\Lambda$  est dense.*

**Démonstration.** Ceci découle du corollaire 10.22, du fait que l'ensemble non errant d'un système dynamique transitif sans point isolé est égal à tout l'espace (voir l'exercice E.49), et du fait qu'une pièce basique qui a un point isolé est réduite à une orbite périodique.  $\square$

## 10.10 Exercices

**Exercice E.58.** Soit  $E$  un espace de Banach. Un endomorphisme linéaire continu  $T$  de  $E$  est dit *linéairement structurellement stable* s'il existe  $\epsilon > 0$  tel que si  $S \in \mathcal{L}(E)$  vérifie  $\|T - S\| \leq \epsilon$ , alors  $T$  et  $S$  sont topologiquement conjugués.

- (1) Montrer que pour toute application lipschitzienne  $\varphi : E \rightarrow E$ , il existe une application lipschitzienne  $\tilde{\varphi} : E \rightarrow E$  bornée, qui coïncide avec  $\varphi$  sur la boule unité de  $E$ , telle que  $\text{Lip}(\tilde{\varphi}) \leq 3 \text{Lip}(\varphi)$ .
- (2) Montrer que l'ensemble des automorphismes linéaires hyperboliques de  $E$  est ouvert dans  $\mathcal{L}(E)$ .
- (3) Montrer qu'un automorphisme linéaire hyperbolique est linéairement structurellement stable.

**Exercice E.59.** Montrer qu'il existe pour  $N \geq 4$  un automorphisme linéaire du tore  $\mathbb{T}^N$ , qui est mélangeant pour la mesure de Haar du tore  $\mathbb{T}^N$ , et qui n'est pas structurellement stable.

**Exercice E.60. (Lemme de pistage linéaire)** Soit  $T$  un automorphisme linéaire continu hyperbolique d'un espace de Banach  $E$ , dont la norme  $\|\cdot\|$  est adaptée à  $T$ , et  $\epsilon > 0$ . Montrer que, pour toute suite  $(y_n)_{n \in \mathbb{Z}}$  dans  $E$  telle que

$$\sup_{n \in \mathbb{Z}} \|y_{n+1} - T(y_n)\| \leq \epsilon,$$

il existe un unique point  $y$  de  $E$  tel que

$$\sup_{n \in \mathbb{Z}} \|y_n - T^n(y)\| \leq \frac{\epsilon}{1 - \text{ch}(T)}.$$

Autrement dit, toute  $\epsilon$ -pseudo-orbite de  $T$  est  $\epsilon'$ -pistée par une unique (vraie) orbite avec  $\epsilon' = \frac{\epsilon}{1 - \text{ch}(T)}$ . On pourra se ramener au cas  $\|T\| < 1$  et considérer  $y = \lim_{n \rightarrow +\infty} T^n(y_{-n})$ .

**Exercice E.61.** Soient  $M$  une variété différentielle lisse et  $k \in \mathbb{N} - \{0\}$ . Rappelons qu'un point fixe  $x_0$  d'un  $C^1$ -difféomorphisme  $\phi$  de  $M$  est hyperbolique si son application tangente  $T_{x_0}\phi \in \mathcal{L}(T_{x_0}M)$  n'a pas de valeur propre (complexe) de module 1. Notons  $\text{Diff}^k(M)$  l'espace de Baire des  $C^k$ -difféomorphismes de  $M$  muni de la convergence uniforme sur les compacts des applications et de leurs applications tangentes jusqu'à l'ordre  $k$ .<sup>233</sup> Nous supposons que  $M$  est compacte dans cet exercice.

(1) Pour tous les  $p, q \in \mathbb{N}$ , construire un  $C^1$ -difféomorphisme du cercle préservant l'orientation ayant exactement  $2p$  points fixes, tous hyperboliques, et, si  $q \geq 1$ , un  $C^1$ -difféomorphisme du cercle préservant l'orientation ayant exactement  $p + q$  points fixes, dont  $p$  hyperboliques et  $q$  non hyperboliques.

(2) Soit  $\phi$  un  $C^1$ -difféomorphisme du cercle préservant l'orientation ayant un nombre fini non nul de points fixes, tous hyperboliques.

• Montrer que  $\phi$  admet un nombre pair de points fixes, notés  $p_1, \dots, p_{2k}$  dans l'ordre trigonométrique.

<sup>233</sup>. Voir les définitions et propriétés de la topologie  $C^k$  avant le corollaire 10.16.

- Montrer qu'il existe  $\epsilon > 0$  assez petit tel que les arcs ouverts de cercle  $A_i$  de longueur  $2\epsilon$  centrés en  $p_i$  sont deux à deux disjoints, et que si  $\phi'$  est  $C^1$ -proche de  $\phi$ , alors  $\phi'$  a exactement un point fixe  $p'_i$  dans chaque arc de cercle de longueur  $2\epsilon$  centré en  $p_i$  et n'a pas de point fixe en dehors des  $A_i$ .

- En fixant un point  $a_i$  de l'arc de cercle entre  $p_i$  et  $p_{i+1}$  qui n'appartient pas à  $A_i \cup A_{i+1}$ , montrer qu'il existe une application continue  $h : \mathbb{S}_1 \rightarrow \mathbb{S}_1$  telle que, pour tout  $i = 1, \dots, 2k$ , nous avons  $h(p_i) = p'_i$ ,  $h(a_i) = a_i$ ,  $h(\phi(a_i)) = \phi'(a_i)$ ,  $h$  est affine sur l'arc de cercle entre  $a_i$  et  $\phi(a_i)$ , et  $h \circ \phi = \phi' \circ h$ .

- En déduire que  $\phi$  est structurellement stable. <sup>234</sup>

(3) Soient  $\phi \in \text{Diff}^k(M)$  et  $U$  un ouvert de  $M$  d'adhérence compacte dans laquelle  $\phi$  admet un et un seul point fixe  $x_\phi$  tels que  $x_\phi \in U$  et l'application tangente de  $\phi$  en ce point soit sans valeur propre égale à 1 (respectivement soit hyperbolique). Montrer qu'il existe un voisinage  $\mathcal{V}$  de  $\phi$  dans  $\text{Diff}^k(M)$  tel que tout  $\phi' \in \mathcal{V}$  admette un et un seul point fixe  $x_{\phi'}$  dans  $\bar{U}$ , et tel que l'application tangente de  $\phi'$  en ce point est sans valeur propre égale à 1 (respectivement sans valeur propre de module égal à 1). On pourra successivement, si 1 n'est pas valeur propre de  $d_{x_\phi}\phi$ , et si  $\phi'$  est  $C^1$ -proche de  $\phi$  :

- se ramener au cas où  $M = \mathbb{R}^N$  et  $x_\phi = 0$  ;
- montrer que l'application  $\tilde{\phi}' : x \mapsto \phi'(x) - x$  est un difféomorphisme d'un voisinage ouvert de 0 sur un voisinage ouvert de 0, et en déduire que  $\phi'$  admet un point fixe  $x_{\phi'}$  ;
- montrer que  $\phi'$  admet au plus un point fixe dans un voisinage de 0 ;
- montrer que si de plus  $d_{x_\phi}\phi$  n'a pas de valeur propre de module égal à 1, et si  $\phi'$  est  $C^1$ -proche de  $\phi$ , alors  $d_{x_{\phi'}}\phi'$  n'a pas non plus de valeur propre de module égal à 1, en considérant les applications  $x \mapsto \phi'(x) - \lambda x$  pour tout  $\lambda \in \mathbb{S}_1$ .

(4) Montrer que le sous-espace  $E_1$  (respectivement  $E_2$ ) de  $\text{Diff}^k(M)$  des  $C^k$ -difféomorphismes de  $M$  ayant un nombre fini de points fixes et en lesquels l'application tangente de  $\phi$  est sans valeur propre égale à 1 (respectivement est hyperbolique) est un ouvert dense. Pour montrer la densité, on pourra successivement

- en appliquant le théorème de Sard <sup>235</sup> à l'application  $f : x \mapsto x - \psi(x)$ , montrer que pour toute application  $\psi : \Omega \rightarrow \mathbb{R}^N$  de classe  $C^1$ , où  $\Omega$  est un ouvert de  $\mathbb{R}^N$ , il existe des  $y$  arbitrairement proches de 0 tels que l'application  $\psi_y : x \mapsto \phi(x) + y$  ait, dans tout compact de  $\Omega$ , un nombre fini de points fixes, en lesquels la différentielle de  $\psi$  n'a pas la valeur propre 1 ;
- montrer que tout point de  $M$  admet un voisinage  $U$  d'adhérence compacte tel que l'ensemble  $E_U$  des  $C^k$ -difféomorphismes de  $M$  qui n'ont qu'un nombre fini de points fixes dans  $\bar{U}$ , ceux-ci appartenant à  $U$ , et en lesquels la différentielle de  $\phi$  n'a pas la valeur propre 1, est un ouvert dense de  $\text{Diff}^k(M)$  ;

---

234. Pour tout  $k \in (\mathbb{N} - \{0\}) \cup \{\infty\}$ , un système dynamique différentiable  $(X, \phi)$  de classe  $C^k$  est dit *structurellement stable* s'il existe un voisinage  $\mathcal{V}$  de  $\phi$  dans  $C^k(X, X)$  pour la topologie  $C^k$  tel que tout élément  $\phi' \in \mathcal{V}$  soit topologiquement conjugué à  $\phi$ . En général, la conjugaison n'est pas de classe  $C^k$ .

235. Si  $f : M \rightarrow N$  est une application  $C^1$  entre deux variétés différentielles lisses, une *valeur critique* de  $f$  est un point  $y \in N$  tel qu'il existe un point  $x \in M$  tel que  $f(x) = y$  et  $T_x f : T_x M \rightarrow T_y N$  ne soit pas surjective. Une *valeur régulière* est un point de  $N$  qui n'est pas valeur critique. Le *théorème de Sard* (voir par exemple le joli petit livre [Mil], ainsi que [Pau2, §6] pour le cas  $m = n$  qui nous intéresse ici) dit que si  $M$  et  $N$  sont deux variétés de dimensions  $m$  et  $n$ , et si  $f : M \rightarrow N$  est une application de classe  $C^k$  où  $k > \max\{0, m - n\}$ , alors l'ensemble des valeurs régulières de  $f$  est dense dans  $N$  (en fait de mesure pleine pour toute mesure sur  $N$  absolument continue par rapport à la mesure de Lebesgue dans chaque carte).

- en déduire la densité cherchée, en utilisant après vérification, dans le cas des points fixes hyperboliques, que pour tout  $\epsilon > 0$ , il existe une fonction  $\varphi_\epsilon : \mathbb{R}^N \rightarrow \mathbb{R}^N$  valant  $x$  en dehors de  $B(0, 2\epsilon)$ , valant  $(1 - \epsilon)x$  sur  $B(0, \epsilon)$  et convergeant en topologie  $C^k$  vers l'identité quand  $\epsilon$  tend vers 0.

(5) Montrer qu'il existe un  $G_\delta$  dense de  $\text{Diff}^k(M)$  constitué d'éléments, ayant, pour tout  $n \in \mathbb{N}$ , un nombre fini de points périodiques de période  $n$ , tous hyperboliques.

(6) Montrer que si  $\phi$  est un  $C^k$ -difféomorphisme structurellement stable de  $M$ , alors  $\phi$  admet, pour tout  $n \in \mathbb{N}$ , un nombre fini de points périodiques de période  $n$ .

**Exercice E.62.** Notons  $\mathcal{M} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$  et  $\phi = \phi_{\mathcal{M}} : x \bmod \mathbb{Z}^2 \mapsto \mathcal{M}x \bmod \mathbb{Z}^2$  le difféomorphisme du tore  $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$  défini par la matrice  $\mathcal{M}$ .

- (1) Montrer que  $0 \in \mathbb{T}^2$  est un point fixe hyperbolique de  $\phi$ , et déterminer les espaces stable  $E_0^s$  et instable  $E_0^u$  dans  $T_0\mathbb{T}^2 = \mathbb{R}^2$ .
- (2) Calculer les variétés stable  $W^s(0)$  et instables  $W^u(0)$  de 0, et montrer qu'elles sont denses dans  $\mathbb{T}^2$ .
- (3) Pour tout  $x \in W^s(0) \cap W^u(0)$  (un tel point différent de 0 est appelé un *point homocline* du point fixe 0), décrire l'orbite complète de  $x$ .

**Exercice E.63.** Soient  $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  le fer à cheval de Smale, et  $\Lambda$  l'attracteur de Smale.

- (1) Calculer les points fixes de  $\phi$ . Pour tout  $x \in \Lambda$ , calculer les espaces stable  $E_x^s$  et instable  $E_x^u$  dans  $T_x\mathbb{R}^2 = \mathbb{R}^2$ . Calculer les variétés stables et instables locales.
- (2) Calculer l'adhérence de l'intersection avec le carré unité  $[0, 1]^2$  d'une variété stable ou instable globale.
- (3) Montrer que  $\Lambda$  vérifie la propriété de produit local.
- (4) Montrer que  $\phi$  est un *difféomorphisme de Kupka-Smale*, c'est-à-dire que tous les points périodiques sont hyperboliques, et que toutes les intersections  $\xi \in W^s(x) \cap W^u(x)$  des variétés stable et instable d'un point périodique  $x$  sont transverses.

## 10.11 Appendice : rappels de théorie spectrale

Nous renvoyons par exemple à [Rud2, LB, Pau6] pour une introduction à la théorie spectrale.

Soient  $E$  un espace de Banach complexe et  $T \in \mathcal{L}(E)$  un endomorphisme linéaire continu de  $E$ . Le *spectre* de  $T$ , noté  $\text{Sp}(T)$  est l'ensemble des  $\lambda \in \mathbb{C}$  tels que  $T - \lambda \text{id}$  soit inversible dans  $\mathcal{L}(E)$  (ou, de manière équivalente par le théorème de Banach, <sup>236</sup> bijectif). L'application  $R_T : \mathbb{C} - \text{Sp}(T) \rightarrow \mathcal{L}(E)$  définie par  $\lambda \mapsto (T - \lambda \text{id})^{-1}$  s'appelle l'*application résolvante* de  $T$ . Le *rayon spectral* de  $T$  est

$$\rho(T) = \sup_{\lambda \in \text{Sp}(T)} |\lambda|$$

(avec la convention usuelle que  $\rho(T) = -\infty$  si  $\text{Sp}(T)$  est vide).

<sup>236.</sup> Le *théorème de Banach* dit que si  $E$  et  $F$  sont deux espaces de Banach réels ou complexes et si  $f : E \rightarrow F$  est une application linéaire continue bijective, alors  $f^{-1} : F \rightarrow E$  est aussi continue.

L'opérateur  $T$  est inversible si et seulement si  $0 \notin \text{Sp}(T)$ , et alors

$$\text{Sp}(T^{-1}) = \text{Sp}(T)^{-1} = \left\{ \frac{1}{\lambda} : \lambda \in \text{Sp}(T) \right\}.$$

En particulier, pour tout  $\kappa_u > 1$ , si  $\text{Sp}(T) \subset \{z \in \mathbb{C} : |z| > \kappa_u\}$ , alors  $T$  est inversible et

$$\rho(T^{-1}) < \frac{1}{\kappa_u} < 1.$$

Nous renvoyons par exemple à [Pau6, §1.3] pour une démonstration du résultat suivant.

**Théorème 10.34.** Soient  $E$  un espace de Banach complexe et  $T \in \mathcal{L}(E)$ .

(i) Le spectre  $\text{Sp}(T)$  de  $T$  est un compact de  $\mathbb{C}$ , et

$$\rho(T) \leq \|T\|.$$

(ii) Si  $\text{Sp}(T)$  est non vide, alors

$$\rho(T) = \lim_{n \rightarrow +\infty} \|T^n\|^{\frac{1}{n}}.$$

(iii) Le spectre  $\text{Sp}(T)$  de  $T$  est non vide si et seulement si  $E \neq \{0\}$ .

(iv) L'application résolvente  $R_T : \mathbb{C} - \text{Sp}(T) \rightarrow \mathcal{L}(E)$  est holomorphe<sup>237</sup> et les éléments de son image commutent deux à deux.  $\square$

Soient  $E$  un espace de Banach réel et  $T \in \mathcal{L}(E)$  un endomorphisme linéaire continu de  $E$ . Notons  $E_{\mathbb{C}}$  l'espace de Banach complexe complexifié de  $E$  et  $T_{\mathbb{C}}$  l'extension de  $T$  à  $E_{\mathbb{C}}$ . Faisons des rappels sur la procédure de complexification des espaces de Banach réels, que le lecteur savant pourra contourner.

### Sur la complexification des espaces de Banach réels

L'espace vectoriel complexe  $E_{\mathbb{C}}$  est l'unique (modulo unique isomorphisme d'espaces vectoriels complexes valant l'identité sur  $E$ ) espace vectoriel complexe dont l'espace vectoriel réel sous-jacent  $(E_{\mathbb{C}})_{\mathbb{R}}$  contient  $E$  et est somme directe (au sens des espaces vectoriels réels)  $(E_{\mathbb{C}})_{\mathbb{R}} = E \oplus iE$ . Tout élément de  $E_{\mathbb{C}}$  s'écrit de manière unique  $x + iy$  avec  $x, y \in E$ . L'espace vectoriel réel  $(E_{\mathbb{C}})_{\mathbb{R}}$  admet une involution  $\mathbb{R}$ -linéaire canonique  $\iota : E_{\mathbb{C}} \rightarrow E_{\mathbb{C}}$ , définie par  $z = x + iy \mapsto \bar{z} = x - iy$  d'ensemble des points fixes  $E$ , appelée la *conjugaison complexe*. Notons que

$$\forall \lambda \in \mathbb{C}, \forall z \in E_{\mathbb{C}}, \quad \iota(\lambda z) = \bar{\lambda} \iota(z).$$

Un sous-espace vectoriel complexe  $F'$  de  $E_{\mathbb{C}}$  est le complexifié  $F_{\mathbb{C}}$  d'un sous-espace vectoriel réel  $F$  de  $E$  si et seulement s'il est stable par  $\iota$ , et alors  $F = F' \cap E$ .

Toute norme  $\|\cdot\|$  sur l'espace vectoriel réel  $E$  s'étend (par une formule surprenamment subtile, mais élémentaire à vérifier) en une norme  $\|\cdot\|_{\mathbb{C}}$  sur l'espace vectoriel complexe  $E_{\mathbb{C}}$  par

$$\forall x, y \in E, \quad \|x + iy\|_{\mathbb{C}} = \max_{\theta \in \mathbb{R}/2\pi\mathbb{Z}} \|\cos \theta x + \sin \theta y\|.$$

Si  $(E, \|\cdot\|)$  est un espace de Banach, alors  $(E_{\mathbb{C}}, \|\cdot\|_{\mathbb{C}})$  est un espace de Banach. Ceci est élémentaire à vérifier, en remarquant que  $\|x + iy\|_{\mathbb{C}} \geq \max\{\|x\|, \|y\|\}$ .

<sup>237</sup>. Notons que l'espace vectoriel  $\mathcal{L}(E)$  des endomorphismes linéaires continus de  $E$ , muni de la norme d'opérateur, est un espace de Banach complexe.

Toute application  $\mathbb{R}$ -linéaire  $T : E \rightarrow E$  s'étend de manière unique en une application  $\mathbb{C}$ -linéaire  $T_{\mathbb{C}}$  de  $E_{\mathbb{C}}$  par  $T_{\mathbb{C}}(x + iy) = T(x) + iT(y)$ . Nous avons les propriétés dites *fonctorielles*  $\text{id}_{\mathbb{C}} = \text{id}$  et  $(T \circ T')_{\mathbb{C}} = T_{\mathbb{C}} \circ T'_{\mathbb{C}}$ . En particulier,  $(T^n)_{\mathbb{C}} = (T_{\mathbb{C}})^n$  pour tout  $n \in \mathbb{N}$  et si  $T$  est inversible, alors  $T_{\mathbb{C}}$  l'est et  $(T_{\mathbb{C}})^{-1} = (T^{-1})_{\mathbb{C}}$ . De même, si  $T_{\mathbb{C}}$  est inversible, alors puisque  $\mathbb{T}_{\mathbb{C}}$ , considérée comme une application  $\mathbb{R}$ -linéaire de  $(E_{\mathbb{C}})_{\mathbb{R}}$  dans  $(E_{\mathbb{C}})_{\mathbb{R}}$ , est diagonale par blocs dans la décomposition en somme directe vectorielle réelle  $(E_{\mathbb{C}})_{\mathbb{R}} = E \oplus iE$ , et vaut  $T$  sur le premier bloc, l'opérateur  $T$  est inversible. L'extension  $T_{\mathbb{C}}$  commute avec la conjugaison complexe :

$$T_{\mathbb{C}} \circ \iota = \iota \circ T_{\mathbb{C}} .$$

Si  $T$  est continue pour une norme  $\| \cdot \|$  sur  $E$ , alors  $T_{\mathbb{C}}$  est continue pour  $\| \cdot \|_{\mathbb{C}}$ , et a même norme d'opérateur :

$$\|T_{\mathbb{C}}\|_{\mathbb{C}} = \|T\| .$$

En effet, nous avons

$$\begin{aligned} \|T_{\mathbb{C}}(x + iy)\|_{\mathbb{C}} &= \|T(x) + iT(y)\|_{\mathbb{C}} = \max_{\theta \in \mathbb{R}/2\pi\mathbb{Z}} \|\cos \theta T(x) + \sin \theta T(y)\| \\ &= \max_{\theta \in \mathbb{R}/2\pi\mathbb{Z}} \|T(\cos \theta x + \sin \theta y)\| \leq \|T\| \|x + iy\|_{\mathbb{C}} , \end{aligned}$$

donc  $\|T_{\mathbb{C}}\|_{\mathbb{C}} \leq \|T\|$  et comme  $(T_{\mathbb{C}})|_E = T$ , l'inégalité inverse est immédiate.

Après ces rappels sur la complexification des espaces de Banach réels, nous définissons le *spectre* de  $T$ , noté  $\text{Sp}(T)$ , par

$$\text{Sp}(T) = \text{Sp}(T_{\mathbb{C}}) . \quad (128)$$

Notons que  $T_{\mathbb{C}} - \bar{\lambda} \text{id} = \iota \circ (T_{\mathbb{C}} - \lambda \text{id}) \circ \iota$  où  $\iota$  est la conjugaison complexe de  $E_{\mathbb{C}}$ . Donc (une propriété bien connue en dimension finie) le spectre de  $T$  est invariant par la conjugaison complexe : pour tout  $\lambda \in \mathbb{C}$ , nous avons  $\lambda \in \text{Sp}(T)$  si et seulement si  $\bar{\lambda} \in \text{Sp}(T)$ , ou autrement dit

$$\text{Sp}(T) = \overline{\text{Sp}(T)} . \quad (129)$$

**Proposition 10.35.** *Soient  $E$  un espace de Banach complexe,  $T \in \mathcal{L}(E)$  et  $\kappa_s \in ]0, 1[$ . Les conditions suivantes sont équivalentes :*

- (1)  $\rho(T) < \kappa_s$ ,
- (2) *il existe une norme  $\| \cdot \|'$  sur  $E$  équivalente à la norme  $\| \cdot \|$  de  $E$  telle que la norme d'opérateur de  $T$  vérifie  $\|T\|' < \kappa_s$ ,*
- (3) *il existe  $\kappa'_s \in ]0, \kappa_s[$  et  $c > 0$  tels que pour tout  $n \in \mathbb{N}$ , nous ayons  $\|T^n\| \leq c (\kappa'_s)^n$  .*

**Démonstration.** Le résultat est immédiat si  $E$  est nul. Supposons donc que  $E$  n'est pas nul. Le fait que (1) implique (3) découle du fait que  $\rho(T) = \lim_{n \rightarrow +\infty} \|T^n\|^{\frac{1}{n}}$  en prenant pour  $\kappa'_s$  n'importe quel élément de  $]\rho(T), \kappa_s[$ .

Le fait que (3) implique (2) se vérifie comme dans la démonstration du lemme 10.1 (3) en posant, pour  $N$  assez grand,

$$\|x\|' = \sum_{i=0}^{N-1} \kappa_s^{-i} \|T^i(x)\| .$$

Le fait que (2) implique (1) découle du fait que  $\lim_{n \rightarrow +\infty} \|T^n\|^{\frac{1}{n}} = \lim_{n \rightarrow +\infty} (\|T^n\|')^{\frac{1}{n}}$  puisque les deux normes sont équivalentes.  $\square$



**Proposition 10.36.** Soient  $E$  un espace de Banach complexe et  $T \in \mathcal{L}(E)$ . Supposons que  $\text{Sp}(T) = \sigma_1 \cup \sigma_2$  avec  $\sigma_1$  et  $\sigma_2$  des compacts disjoints, tels qu'il existe une courbe de Jordan lisse  $\gamma : \mathbb{S}_1 \rightarrow \mathbb{C}$  contenant  $\sigma_1$  dans son intérieur et  $\sigma_2$  dans son extérieur. Alors il existe deux sous-espaces vectoriels fermés  $E_1$  et  $E_2$  de  $E$  tels que

$$E = E_1 \oplus E_2, \quad T(E_1) \subset E_1, \quad T(E_2) \subset E_2, \quad \text{Sp}(T|_{E_1}) = \sigma_1, \quad \text{Sp}(T|_{E_2}) = \sigma_2 .$$

**Démonstration.** Soit  $\gamma$  une courbe de Jordan lisse orientée dans le sens trigonométrique, contenant  $\sigma_1$  dans la composante connexe bornée de  $\mathbb{C} - \gamma$  et  $\sigma_2$  dans la composante connexe non bornée de  $\mathbb{C} - \gamma$ . En déformant un peu  $\gamma$  par compacité de  $\sigma_2$ , soit  $\gamma'$  une courbe de Jordan lisse orientée dans le sens trigonométrique, contenant  $\gamma$  dans la composante connexe bornée de  $\mathbb{C} - \gamma'$  et  $\sigma_2$  dans la composante connexe non bornée de  $\mathbb{C} - \gamma'$ . Par le théorème des résidus, puisque  $\gamma$  est dans la composante connexe bornée du complémentaire de  $\gamma'$ , nous avons

$$\forall \lambda' \in \gamma', \quad \frac{1}{2i\pi} \int_{\gamma} \frac{d\lambda}{\lambda - \lambda'} = 0 \quad \text{et} \quad \forall \lambda \in \gamma, \quad \frac{1}{2i\pi} \int_{\gamma'} \frac{d\lambda'}{\lambda - \lambda'} = 1 . \quad (130)$$

Puisque  $\gamma$  et  $\gamma'$  sont contenus dans  $\mathbb{C} - \text{Sp}(T)$  et par le théorème 10.34 (iv), l'application  $\lambda \mapsto R_\lambda(T)$  est analytique complexe sur un voisinage de l'anneau de bords  $\gamma$  et  $\gamma'$ . Le théorème des résidus dit donc que

$$\frac{1}{2i\pi} \int_{\gamma} R_T(\lambda) d\lambda = \frac{1}{2i\pi} \int_{\gamma'} R_T(\lambda) d\lambda .$$

Notons  $P_1 \in \mathcal{L}(E)$  cet opérateur linéaire continu et  $P_2 = \text{id} - P_1$ . Montrons que  $P_1$  est un projecteur linéaire continu, c'est-à-dire que  $P_1^2 = P_1$ . Ceci montrera que  $P_2$  est aussi un projecteur linéaire continu (car alors  $(\text{id} - P_1)^2 = \text{id} - 2P_1 - P_1^2 = \text{id} - P_1$ ), que  $E_1 = P_1(E)$  et  $E_2 = P_2(E)$  sont fermés (car  $E_1 = \ker P_2$  et  $E_2 = \ker P_1$ ) et que  $E$  est somme directe de  $E_1$  et de  $E_2$  (car  $\text{id} = P_1 + P_2$  et  $\ker P_1 \cap P_1(E) = \{0\}$ ).

Nous avons, pour tous les  $\lambda, \lambda' \in \mathbb{C} - \text{Sp}(T)$ ,

$$R_T(\lambda) = R_T(\lambda)(T - \lambda \text{id} + (\lambda - \lambda') \text{id})R_T(\lambda') = R_T(\lambda') + (\lambda - \lambda')R_T(\lambda)R_T(\lambda') . \quad (131)$$

D'où, en utilisant le théorème de Fubini et les formules (131) et (130),

$$\begin{aligned} P_1^2 &= \frac{1}{(2i\pi)^2} \int_{\gamma'} \int_{\gamma} R_T(\lambda)R_T(\lambda') d\lambda d\lambda' = \frac{1}{(2i\pi)^2} \int_{\gamma'} \int_{\gamma} \frac{R_T(\lambda) - R_T(\lambda')}{\lambda - \lambda'} d\lambda d\lambda' \\ &= \frac{1}{2i\pi} \int_{\gamma} R_T(\lambda) \frac{1}{2i\pi} \int_{\gamma'} \frac{d\lambda'}{\lambda - \lambda'} d\lambda + \frac{1}{2i\pi} \int_{\gamma'} R_T(\lambda') \frac{1}{2i\pi} \int_{\gamma} \frac{d\lambda}{\lambda - \lambda'} d\lambda' = P_1 . \end{aligned}$$

Puisque  $T$  commute avec  $R_T(\lambda)$  pour tout  $\lambda \in \mathbb{C} - \text{Sp}(T)$ , il commute avec  $P_1$  par intégration et avec  $P_2 = \text{id} - P_1$ . Donc  $T$  préserve  $E_1 = P_1(E)$  et  $E_2 = P_2(E)$ .

Montrons que

$$\text{Sp}(T) = \text{Sp}(T|_{E_1}) \cup \text{Sp}(T|_{E_2}) . \quad (132)$$

En effet, pour tout  $\lambda \in \mathbb{C}$ , puisque  $E = E_1 \oplus E_2$  et comme  $E_1$  et  $E_2$  sont fermés et invariants par  $T - \lambda \text{id}$ , l'endomorphisme linéaire continu  $T - \lambda \text{id}$  de  $E$  est inversible si et seulement si ses restrictions à  $E_1$  et à  $E_2$  le sont.

[Voici une autre méthode, plus longue. Pour tout  $\lambda \in \mathbb{C} - \text{Sp}(T)$ , en intégrant sur  $\lambda' \in \gamma'$  la relation de commutation

$$R_T(\lambda)R_T(\lambda') = R_T(\lambda')R_T(\lambda) ,$$

nous avons  $P_1R_T(\lambda) = R_T(\lambda)P_1$ . En particulier  $R_T(\lambda)$  préserve  $E_1 = P_1(E)$ , et

$$R_T(\lambda)|_{E_1} (T|_{E_1} - \lambda \text{id}) = (R_T(\lambda)(T - \lambda \text{id})|_{E_1} = \text{id}_{E_1} = (T|_{E_1} - \lambda \text{id})R_T(\lambda)|_{E_1} .$$

Donc  $\lambda \in \mathbb{C} - \text{Sp}(T|_{E_1})$ . De même,  $\lambda \in \mathbb{C} - \text{Sp}(T|_{E_2})$ .

Réciproquement, si  $\lambda \in (\mathbb{C} - \text{Sp}(T|_{E_1})) \cap (\mathbb{C} - \text{Sp}(T|_{E_2}))$ , nous avons

$$(R_{T|_{E_1}}(\lambda) \oplus R_{T|_{E_2}}(\lambda))(T - \lambda \text{id}) = \text{id}_E = (T - \lambda \text{id})(R_{T|_{E_1}}(\lambda) \oplus R_{T|_{E_2}}(\lambda)) ,$$

donc  $\lambda \in \mathbb{C} - \text{Sp}(T)$ . Ceci remonte la formule (132) par passage aux complémentaires.]

Notons  $\bar{\Omega}$  l'adhérence de la composante connexe bornée de  $\mathbb{C} - \gamma$ . Elle contient  $\sigma_1$  dans son intérieur et son complémentaire contient  $\sigma_2$ . Soit  $\lambda' \in \mathbb{C} - \bar{\Omega}$ .

Remarquons que pour tout  $\lambda \in \gamma$ , nous avons

$$(T - \lambda' \text{id})R_T(\lambda) = ((\lambda - \lambda') \text{id} + (T - \lambda \text{id}))R_T(\lambda) = (\lambda - \lambda')R_T(\lambda) + \text{id} .$$

Donc en divisant par  $\lambda - \lambda'$  et par intégration sur  $\lambda \in \gamma$ , en notant  $T' = \frac{1}{2i\pi} \int_{\gamma} \frac{R_T(\lambda)}{\lambda - \lambda'} d\lambda$ , qui préserve  $E_1$ , nous avons, par la formule (130),

$$(T - \lambda' \text{id})T' = \frac{1}{2i\pi} \int_{\gamma} R_T(\lambda) d\lambda + \int_{\gamma} \frac{d\lambda}{\lambda - \lambda'} \text{id} = P_1 .$$

En particulier,  $(T|_{E_1} - \lambda' \text{id})T'|_{E_1} = T'|_{E_1} (T|_{E_1} - \lambda' \text{id}) = \text{id}_{E_1}$ . Donc  $\lambda' \in \mathbb{C} - \text{Sp}(T|_{E_1})$ . Par conséquent,  $\text{Sp}(T|_{E_1}) \subset \bar{\Omega}$ . Un argument similaire montre que  $\text{Sp}(T|_{E_2})$  est contenu dans l'adhérence du complémentaire de la composante connexe bornée de  $\mathbb{C} - \gamma'$ .

Par la formule (132), nous avons donc  $\text{Sp}(T|_{E_1}) = \bar{\Omega} \cap \text{Sp}(T) = \sigma_1$ . De même, nous avons  $\text{Sp}(T|_{E_2}) = \sigma_2$ .  $\square$

**Corollaire 10.37.** *Soient  $E$  un espace de Banach réel et  $T \in \mathcal{L}(E)$ . Supposons que nous ayons  $\text{Sp}(T) = \sigma_1 \cup \sigma_2$  avec  $\sigma_1$  et  $\sigma_2$  des compacts disjoints de  $\mathbb{C}$ , tels qu'il existe une courbe de Jordan lisse  $\gamma : \mathbb{S}_1 \rightarrow \mathbb{C}$  contenant  $\sigma_1$  dans son intérieur et  $\sigma_2$  dans son extérieur, et commutant avec la conjugaison complexe, c'est-à-dire vérifiant  $\gamma(e^{-i\theta}) = \overline{\gamma(e^{i\theta})}$  pour tout  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ . Alors il existe deux sous-espaces vectoriels fermés  $E_1$  et  $E_2$  de  $E$  tels que*

$$E = E_1 \oplus E_2, \quad T(E_1) \subset E_1, \quad T(E_2) \subset E_2, \quad \text{Sp}(T|_{E_1}) = \sigma_1, \quad \text{Sp}(T|_{E_2}) = \sigma_2 .$$

**Démonstration.** Nous identifions  $E$  avec son image dans  $E_{\mathbb{C}} = E \oplus iE$  par l'application  $x \mapsto x + i0$ . Puisque  $\text{Sp}(T) = \text{Sp}(T_{\mathbb{C}})$ , par la proposition 10.36, il existe des sous-espaces vectoriels fermés  $E_{1,\mathbb{C}}$  et  $E_{2,\mathbb{C}}$  de  $E_{\mathbb{C}}$  tels que

$$E_{\mathbb{C}} = E_{1,\mathbb{C}} \oplus E_{2,\mathbb{C}}, \quad T_{\mathbb{C}}(E_{1,\mathbb{C}}) \subset E_{1,\mathbb{C}}, \quad T_{\mathbb{C}}(E_{2,\mathbb{C}}) \subset E_{2,\mathbb{C}} ,$$

$$\text{Sp}(T_{\mathbb{C}}|_{E_{1,\mathbb{C}}}) = \sigma_1, \quad \text{Sp}(T_{\mathbb{C}}|_{E_{2,\mathbb{C}}}) = \sigma_2 .$$

De plus, la démonstration de la proposition 10.36 dit que  $E_{1,\mathbb{C}} = P_1(E_{\mathbb{C}})$  où

$$P_1 = \frac{1}{2i\pi} \int_{\gamma} R_{T_{\mathbb{C}}}(\lambda) d\lambda .$$

Nous avons vu que si  $\iota$  est la conjugaison complexe de  $E_{\mathbb{C}}$ , alors  $(T_{\mathbb{C}} - \lambda \text{id}) \circ \iota = \iota \circ (T_{\mathbb{C}} - \bar{\lambda} \text{id})$  pour tout  $\lambda \in \mathbb{C}$ . Rappelons (voir la formule (129)) que le spectre de  $T$  est invariant par conjugaison complexe. Donc pour tout  $\lambda \in \mathbb{C} - \text{Sp}(T)$ , nous avons  $R_{T_{\mathbb{C}}}(\lambda) \circ \iota = \iota \circ R_{T_{\mathbb{C}}}(\bar{\lambda})$ .

Par intégration sur  $\gamma$  et changement de variable  $e^{i\theta} \mapsto e^{-i\theta}$  (ce qui est possible par l'hypothèse sur la courbe  $\gamma$ ), nous avons donc  $P_1 \circ \iota = \iota \circ P_1$ . Donc  $E_{1,\mathbb{C}} = P_1(E_{\mathbb{C}})$  et  $E_{2,\mathbb{C}} = (\text{id} - P_1)(E_{\mathbb{C}})$  sont invariants par  $\iota$ . Donc  $E_1 = E_{1,\mathbb{C}} \cap E$  et  $E_2 = E_{2,\mathbb{C}} \cap E$  sont des sous-espaces vectoriels réels fermés de  $E$  tels que  $E = E_1 \oplus E_2$ . Ils sont invariants par  $T$ , et comme  $(T|_{E_1})_{\mathbb{C}} = (T_{\mathbb{C}})|_{E_{1,\mathbb{C}}}$  et  $(T|_{E_2})_{\mathbb{C}} = (T_{\mathbb{C}})|_{E_{2,\mathbb{C}}}$ , le résultat en découle.  $\square$

**Démonstration de la proposition 10.2.** Comme déjà dit, la seconde assertion implique la première. Soient  $0 < \kappa_s < 1 < \kappa_u$ .

Soit  $T \in \mathcal{L}(E)$  un élément  $(\kappa_s, \kappa_u)$ -hyperbolique. Soient  $E^s$  et  $E^u$  les sous-espaces vectoriels fermés de  $E$  comme dans le début de la partie 10.1. Comme  $E = E^s \oplus E^u$ , et puisque  $E^s$  et  $E^u$  sont fermés et invariants par  $T$  (voir la démonstration de la formule (132)), nous avons  $\text{Sp}(T) = \text{Sp}(T|_{E^s}) \cup \text{Sp}(T|_{E^u})$ . La proposition 10.35 implique que

$$\text{Sp}(T|_{E^s}) \subset \{z \in \mathbb{C} : |z| < \kappa_s\}$$

et

$$(\text{Sp}(T|_{E^u}))^{-1} = \text{Sp}((T|_{E^u})^{-1}) \subset \{z \in \mathbb{C} : |z| < \kappa_u^{-1}\}.$$

Donc  $\text{Sp}(T)$  ne rencontre pas l'anneau  $\{z \in \mathbb{C} : \kappa_s \leq |z| \leq \kappa_u\}$ .

Réciproquement, soit  $T \in \mathcal{L}(E)$  tels que  $\text{Sp}(T) \cap \{z \in \mathbb{C} : \kappa_s \leq |z| \leq \kappa_u\} = \emptyset$ . Soit  $\sigma_1 = \text{Sp}(T) \cap \{z \in \mathbb{C} : |z| \leq 1\}$ ,  $\sigma_2 = \text{Sp}(T) \cap \{z \in \mathbb{C} : |z| \geq 1\}$ , qui sont deux fermés disjoints de réunion  $\text{Sp}(T)$ . La courbe de Jordan  $\gamma : \mathbb{S}_1 \rightarrow \mathbb{C}$  définie par  $e^{i\theta} \mapsto e^{i\theta}$  contient  $\sigma_1$  dans son intérieur et  $\sigma_2$  dans son extérieur, et est invariante par conjugaison. Donc par le corollaire 10.37, l'espace  $E$  est somme directe de deux sous-espaces vectoriels fermés (donc de Banach)  $E_1$  et  $E_2$  préservés par  $T$  tels que  $\text{Sp}(T|_{E_1}) = \sigma_1$  et  $\text{Sp}(T|_{E_2}) = \sigma_2$ . En particulier,  $0 \notin \text{Sp}(T|_{E_2})$ , donc  $(T|_{E_2})_{\mathbb{C}}$  est inversible, donc  $T|_{E_2} : E_2 \rightarrow E_2$  est inversible, et  $\text{Sp}((T|_{E_2})^{-1}) \subset \{z \in \mathbb{C} : |z| \leq \frac{1}{\kappa_u}\}$ . En appliquant la proposition 10.35 à  $T|_{E_1}$  et à  $(T|_{E_2})^{-1}$ , le résultat en découle.  $\square$

## 10.12 Indications pour la résolution des exercices

**Correction de l'exercice E.59.** Notons  $P$  le polynôme  $X^4 + 2X^3 + X^2 + 2X + 1$ , et

$$M = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \end{pmatrix} \text{ sa matrice compagnon. Le polynôme caractéristique } \det(M - \lambda \text{id})$$

de  $M$  est, par récurrence en développant par rapport à la première ligne, égal à  $(-1)^4 P$ . La matrice  $M$  est inversible (car  $P(0) \neq 0$ ) et elle est à coefficients entiers. Les valeurs propres de  $M$  sont les racines de  $P$ , qui sont

$$\frac{1}{2} \left( -\sqrt{2} - 1 - \sqrt{2\sqrt{2} - 1} \right) \simeq -1,883, \quad \frac{1}{2} \left( -\sqrt{2} - 1 + \sqrt{2\sqrt{2} - 1} \right) \simeq -0,531,$$

$$\frac{1}{2} \left( \sqrt{2} - 1 - i\sqrt{2\sqrt{2} + 1} \right), \quad \frac{1}{2} \left( \sqrt{2} - 1 + i\sqrt{2\sqrt{2} + 1} \right).$$

Il est possible de montrer qu'aucune racine de  $P$  n'est une racine de l'unité. Donc la transformation  $\phi_M$  du tore définie par  $M$  est mélangeante pour la mesure de Haar du tore,

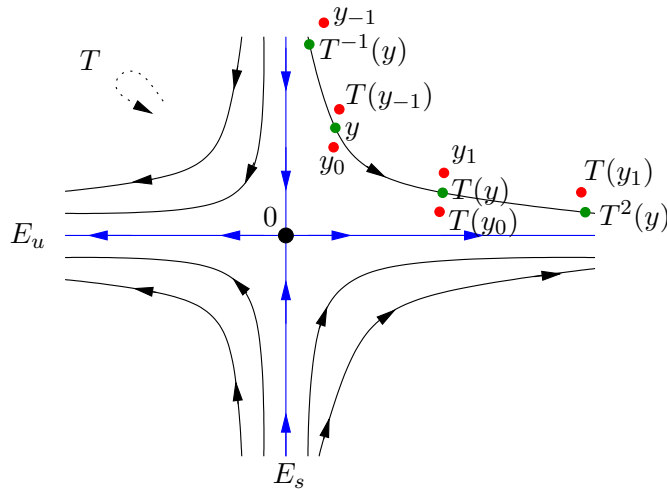
par le théorème 4.10. Il est possible de montrer que  $P$  admet une valeur propre de module 1 (les deux dernières ci-dessus, qui sont conjuguées), et qu'il est possible de perturber  $\phi$  de manière lisse et arbitrairement petite sans être topologiquement conjugué, en perturbant la rotation induite par  $M$  dans le plan vectoriel réel invariant par  $M$  correspondant aux deux dernières valeurs propres, par l'exemple donné dans la remarque suivant le théorème de Grobman-Hartman 10.6.

Voici une méthode pour trouver d'autres exemples de construction de polynôme entier (afin que sa matrice compagnon  $M$  soit entière)  $P$  de degré 4 ayant deux racines complexes conjuguées de module 1 qui ne sont pas des racines de l'unité, une racine réelle  $a > 1$  et une racine réelle  $\frac{1}{a}$  (de sorte que sa matrice compagnon soit de déterminant 1, donc inversible sur  $\mathbb{Z}$ ). Pour tout tel polynôme  $P$ , l'automorphisme linéaire  $\phi_M$  du tore  $\mathbb{T}^4$  associé à la matrice compagnon  $M$  de  $P$  fournit une solution de l'exercice E.59.

Si  $P$  est un tel polynôme, c'est nécessairement un palindrome, c'est-à-dire un polynôme tel que  $x \in \mathbb{C}$  soit une racine si et seulement si  $\frac{1}{x}$  est une racine, dont le coefficient constant (le produit des racines) et donc le coefficient dominant sont égaux à 1. Il existe alors un polynôme  $Q$  quadratique unitaire à coefficients entiers tel que  $P(X) = X^2Q(X + \frac{1}{X})$ . Si  $Q$  admet une racine réelle  $\lambda > 2$ , alors les deux solutions réelles de l'équation  $X + \frac{1}{X} = \lambda$  sont de la forme  $a > 1$  et  $\frac{1}{a}$ . Si  $Q$  admet une racine réelle  $\mu$  dans  $] - 2, 2[$ , alors les deux nombres complexes conjugués  $z$  et  $\bar{z}$  de module 1 dont la partie réelle est égale à  $\frac{\mu}{2}$  sont des racines de  $P$ , car  $z + \frac{1}{z} = z + \bar{z} = 2\frac{\mu}{2} = \mu$ .

Nous nous ramenons ainsi à déterminer des polynômes quadratiques unitaires entiers ayant une valeur propre  $> 2$  et une dans  $] - 2, 2[$ , ce qui se résoud algorithmiquement par les formules du binôme.

**Correction de l'exercice E.60.** En écrivant  $y_n = (y_n^s, y_n^u)$  dans la décomposition en sous-espaces stable et instable  $E = E^s \oplus E^u$  de  $T$ , qui est  $T$ -invariante, et puisque la norme est adaptée, nous avons  $\sup_{n \in \mathbb{Z}} \|y_{n+1}^s - T(y_n^s)\| \leq \epsilon$  et  $\sup_{n \in \mathbb{Z}} \|y_{n+1}^u - T(y_n^u)\| \leq \epsilon$ . En se restreignant à  $E^s$  et  $E^u$  et en remplaçant  $T$  par  $T^{-1}$  sur  $E^u$ , nous pouvons donc supposer que  $E = E^s$  et  $\|T\| < 1$ .



Comme

$$\|T^{n+1}(y_{-(n+1)}) - T^n(y_{-n})\| \leq \|T\|^n \|T(y_{-n-1}) - y_{-n}\| \leq \epsilon \|T\|^n,$$

la suite  $(T^n(y_{-n}))_{n \in \mathbb{N}}$  est de Cauchy dans l'espace complet  $E$ , donc converge vers un élément  $y$  de  $E$ .

Pour tout  $k$  dans  $\mathbb{Z}$ , si  $n \geq 1 - k$ , nous avons

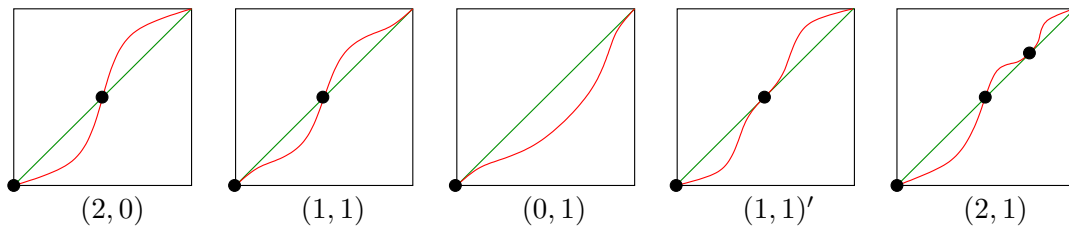
$$y_k - T^k(y) = \sum_{i=0}^{n+k-1} (T^i(y_{k-i}) - T^{i+1}(y_{k-i-1})) + T^{k+n}y_{-n} - T^k y .$$

Donc

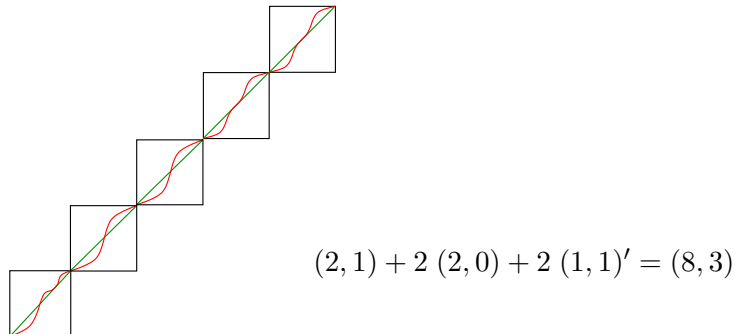
$$\|y_k - T^k(y)\| \leq \sum_{i=0}^{n+k-1} \|T\|^i \|y_{k-i} - T^1(y_{k-i-1})\| + \|T^k\| \|T^n(y_{-n}) - y\| .$$

En faisant tendre  $n$  vers  $+\infty$ , nous avons donc  $\|y_k - T^k(y)\| \leq \frac{\epsilon}{1-\|T\|}$ .

**Correction de l'exercice E.61.** (1) Une rotation d'angle non nul (modulo  $2\pi$ ) est un  $C^1$ -difféomorphisme préservant l'orientation sans point fixe. Un  $C^1$ -difféomorphisme  $\phi$  du cercle préservant l'orientation et ayant au moins un point fixe est, modulo identifications, un  $C^1$ -difféomorphisme strictement croissant de  $[0, 1]$  ayant mêmes dérivées en 0 et 1. Un point fixe de  $\phi$  correspond à un point d'intersection du graphe de  $\phi$  avec la diagonale du carré. Il est hyperbolique si et seulement si la dérivée de  $\phi$  en ce point est différente de 1. Les exemples sont obtenus en mettant bout à bout des blocs  $(p, q)$  avec  $p$  points fixes hyperboliques (de dérivée 2 ou  $\frac{1}{2}$ ) et  $q$  points fixes non hyperboliques pour  $(p, q)$  valant  $(2, 0)$ ,  $(1, 1)$  (avec deux sous-types  $(1, 1)$  et  $(1, 1)'$ ),  $(0, 1)$ , et  $(2, 1)$ .

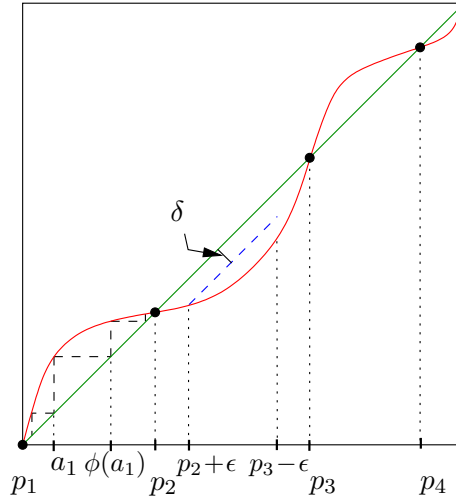


Si  $q = 0$  (auquel cas  $p$  est supposé pair), on prend  $\frac{p}{2}$  blocs  $(2, 0)$ . Supposons  $q \geq 1$ . Si  $p \leq q$ , on prend  $p$  blocs  $(1, 1)$  et  $q - p$  blocs  $(0, 1)$ . Si  $p > q$  et  $p - q$  est pair, on prend  $q$  blocs  $(1, 1)'$  et  $\frac{p-q}{2}$  blocs  $(2, 0)$ . Enfin, si  $p > q$  et  $p - q$  est impair, on prend  $q - 1$  blocs  $(1, 1)'$ ,  $\frac{p-q-1}{2}$  blocs  $(2, 0)$  et 1 bloc  $(2, 1)$ .



(2) • Par le théorème des valeurs intermédiaires, les dérivées de  $\phi$  aux points fixes de  $\phi$  doivent être, lorsque l'on parcourt le cercle dans le sens trigonométrique, alternativement strictement supérieure à 1 et strictement inférieure à 1. Donc le nombre de points fixes de  $\phi$  est pair.

• Fixons  $\epsilon, \epsilon' > 0$  assez petits tels que les arcs ouverts de cercle  $A_i$  de longueur  $2\epsilon$  centrés en  $p_i$  sont deux à deux disjoints et tels que la dérivée de  $\phi$  sur l'adhérence de  $A_i$  soit à distance au moins  $\epsilon'$  de 1. En dehors de ces arcs, la distance du graphe de  $\phi$  à la diagonale est supérieure à une constante  $\delta > 0$ . Si  $\phi'$  est  $C^1$ -proche de  $\phi$ , alors sa dérivée dans l'adhérence de chaque  $A_i$  est à distance au moins  $\frac{\epsilon'}{2}$  de 1, et la distance du graphe de  $\phi'$  à la diagonale est supérieure à  $\frac{\delta}{2} > 0$ . En particulier,  $\phi'$  n'a pas de point fixe en dehors des  $A_i$ .



Par le théorème des valeurs intermédiaires, le graphe de  $\phi'$  rencontre la diagonale en au moins un point de  $A_i$ , et au plus en un point car sinon, encore par le théorème des valeurs intermédiaires, la dérivée de  $\phi'$  serait égale à 1 en au moins un point de  $A_i$ .

Puisque  $p'_i \in ]p_i - \epsilon, p_i + \epsilon[$ ,  $p'_{i+1} \in ]p_{i+1} - \epsilon, p_{i+1} + \epsilon[$ , et  $a_i \in ]p_i + \epsilon, p_{i+1} + \epsilon[$ , nous avons  $a_i \in ]p'_i + \epsilon, p'_{i+1} + \epsilon[$ . La restriction de  $\phi$  à l'arc de cercle  $]p_i, p_{i+1}[$  est un difféomorphisme strictement croissant dans lui-même. Si  $\phi(a_i) > a_i$ , alors pour tout  $x \in ]p_i, p_{i+1}[$ , la suite  $(\phi^n(x))_{n \in \mathbb{Z}}$  est une suite strictement croissante dans  $]p_i, p_{i+1}[$ , qui converge vers  $p_{i+1}$  quand  $n \rightarrow +\infty$  et vers  $p_i$  quand  $n \rightarrow -\infty$ . Donc  $\phi(a_i) > a_i$  est équivalent au fait que  $\phi(p_i) > 1$ , donc à  $\phi'(p'_i) > 1$ , donc à  $\phi'(a_i) > a_i$ . En particulier  $]p_i, p_{i+1}[$  est la réunion disjointe des arcs de cercle  $\phi^n([a_i, \phi(a_i)[) = ]\phi^n(a_i), \phi^{n+1}(a_i)[$  pour  $n \in \mathbb{Z}$ . De même,  $]p'_i, p'_{i+1}[$  est la réunion disjointe des arcs de cercle  $](\phi')^n(a_i), (\phi')^{n+1}(a_i)[$  pour  $n \in \mathbb{Z}$ . Si nous posons  $h(p_i) = p'_i$ ,  $h(a_i) = a_i$ ,  $h(\phi(a_i)) = \phi'(a_i)$ , et  $h$  affine de  $[a_i, \phi(a_i)[$  dans  $[a_i, \phi'(a_i)[$ , nous pouvons donc étendre  $h : [a_i, \phi(a_i)[ \rightarrow [a_i, \phi'(a_i)[$  de manière unique en une application croissante continue  $h$  de  $]p_i, p_{i+1}[$  dans  $]p'_i, p'_{i+1}[$  pour que, sur cet intervalle  $]p_i, p_{i+1}[$ , nous ayons la relation de conjugaison  $h \circ \phi = \phi' \circ h$  : pour tous les  $n \in \mathbb{Z}$  et  $t \in ]\phi^n(a_i), \phi^{n+1}(a_i)[$ , nous posons

$$h(t) = (\phi')^n(h(\phi^{-n}(t))),$$

ce qui est bien défini car  $\phi^{-n}(t)$  appartient à  $[a_i, \phi(a_i)[$ .

En procédant sur chaque intervalle  $]p_i, p_{i+1}[$  pour  $i = 1, \dots, 2k$  (en procédant de manière similaire si  $\phi(a_i) < a_i$ ), nous définissons donc une application continue bijective  $h$  du

cercle dans lui-même, qui est donc un homéomorphisme par compacité, qui conjugue  $\phi$  et  $\phi'$ . Ceci montre que  $\phi$  est structurellement stable.

(3) • Pour tout voisinage ouvert  $U'$  de  $x_\phi$  d'adhérence compacte et contenue dans  $U$ , si  $\phi'$  est uniformément proche de  $\phi$ , alors comme l'application continue  $x \mapsto d(\phi(x), x)$  a un minimum  $\delta$  strictement positif sur le compact  $\bar{U} - U'$ , l'application  $x \mapsto d(\phi'(x), x)$  sera minorée par  $\frac{\delta}{2}$ , donc l'application  $\phi'$  n'aura pas de point fixe dans  $\bar{U} - U'$ . En prenant une carte locale, nous pouvons supposer que  $M = \mathbb{R}^N$  et  $x_T = 0$  et montrer qu'il existe un voisinage  $U'$  de 0 tel que si  $\phi'$  est  $C^1$ -proche de  $\phi$ , alors  $\phi'$  admet un et un seul point fixe dans  $U'$ , avec les mêmes conditions sur les valeurs propres.

- Puisque 1 n'est pas valeur propre de  $d_0\phi$ , la différentielle de  $\tilde{\phi} : x \mapsto \phi(x) - x$  en 0 est inversible. Donc par le théorème d'inversion locale,<sup>238</sup> quitte à restreindre  $U'$ , l'application  $\tilde{\phi}' : x \mapsto \phi'(x) - x$  est un difféomorphisme d'un voisinage de 0 dans un voisinage de 0. En particulier, puisque 0 appartient à l'image de  $\tilde{\phi}'$ , l'application  $\phi'$  admet un point fixe  $x_{\phi'}$ .

- Montrons que quitte à restreindre  $U'$ , l'application  $\phi'$  admet au plus un point fixe. Sinon, il existerait une suite  $(\phi_i)_{i \in \mathbb{N}}$  qui converge pour la topologie  $C^k$  vers  $\phi$ , de sorte que  $\phi_i$  admette deux points fixes distincts  $x_{\phi_i}$  et  $y_{\phi_i}$  qui convergent vers 0. Mais alors quitte à extraire, par un développement limité à l'ordre 1 de  $\phi_i$  en 0, la suite de vecteurs  $v_i = \frac{x_{\phi_i} - y_{\phi_i}}{\|x_{\phi_i} - y_{\phi_i}\|}$  convergerait vers un vecteur non nul fixe par la différentielle de  $\phi$  en 0, ce qui n'existe pas.

- Prolongeons les applications sur  $\mathbb{R}^N$  à  $\mathbb{C}^N$  de manière constante sur les droites verticales, ce qui préserve les caractères  $C^1$  et  $C^1$ -proche. Par compacité de  $\mathbb{S}_1$ , un raisonnement analogue montre que, quitte à réduire  $U'$ , si l'application  $d_{x_\phi}\phi$  n'a pas de valeur propre de module égal à 1, et si  $\phi'$  est  $C^1$ -proche de  $\phi$ , alors, pour tout  $\lambda \in \mathbb{S}_1$ , l'application  $x \mapsto \phi'(x) - \lambda x$  est de différentielle non nulle dans les directions réelles, donc que  $d_{x_\phi}\phi'$  n'a pas de valeur propre de module égal à 1.

(4) En travaillant sur des voisinages disjoints autour de chaque point fixe, la question (3) montre que  $E_1$  et  $E_2$  sont ouverts dans  $\text{Diff}^k(M)$ .

- Remarquons que  $x$  est un point fixe de  $\psi_y$  si et seulement si  $x \in f^{-1}(y)$ . Par le théorème de Sard, qui s'applique en régularité  $C^1$  car  $\Omega$  et  $\mathbb{R}^N$  ont la même dimension, l'application  $f$  a un ensemble dense de valeurs régulières. Si  $y$  est une valeur régulière de  $f$ , puisqu'un endomorphisme linéaire surjectif est bijectif, et par le théorème d'inversion locale, l'application  $f$  est un difféomorphisme local en tout point de  $f^{-1}(y)$ , et en particulier  $f^{-1}(y)$  est discret, donc son intersection avec tout compact est finie. Pour tout  $x \in f^{-1}(y)$ , comme  $d_x\psi - \text{id} = -d_x f$  est bijectif, nous avons donc que  $d_x\psi$  est sans valeur propre égale à 1. Enfin, si  $y$  est proche de 0, alors l'application  $\psi_y$  est  $C^1$ -proche de l'application  $\psi$ , puisqu'elles ne diffèrent que d'une petite constante.

- L'ouverture de  $E_U$  découle de la question (3). Pour la densité, il suffit d'appliquer en carte locale le point précédent, avec des arguments usuels de fonction plateau, en utilisant le fait que  $\text{Diff}^k(M)$  est ouvert dans  $C^k(M, M)$ .

- Par compacité de  $M$ , nous pouvons recouvrir  $M$  par un nombre fini  $U_1, \dots, U_k$  de voisinages vérifiant la propriété précédente. La densité de l'ensemble des  $C^k$ -difféomorphismes

---

238. La version uniforme (voir par exemple la démonstration de [Car]) du *théorème d'inversion locale* dit que pour tout  $k \in (\mathbb{N} - \{0\}) \cup \{\infty\}$ , si  $\Omega$  est un ouvert de  $\mathbb{R}^N$ , si  $x_0 \in \Omega$ , et si  $f : \Omega \rightarrow \mathbb{R}^N$  est une application de classe  $C^k$  telle que l'application linéaire  $d_{x_0}f$  soit inversible, alors il existe des voisinages ouverts  $U$  de  $x_0$  et  $V$  de  $f(x_0)$  tels que si  $g$  est  $C^k$ -proche de  $f$ , alors  $g$  est un difféomorphisme de  $U$  sur son image et  $V$  est contenu dans  $g(U)$ .

de  $M$  qui n'ont qu'un nombre fini de points fixes dans  $M$ , et en lesquels la différentielle de  $\phi$  n'a pas la valeur propre 1, découle alors du fait qu'une intersection finie  $E_{U_1} \cap \dots \cap E_{U_k}$  d'ouverts denses est encore ouverte et dense.

Pour passer au cas des points fixes hyperboliques, il suffit alors d'approcher, en topologie  $C^k$ , un  $C^k$ -difféomorphisme  $\phi$  de classe  $C^k$ , ayant un nombre fini de points fixes, et en lesquels la différentielle de  $\phi$  n'a pas la valeur propre 1, par une application  $C^k$  qui coïncide avec  $\phi$  en dehors d'un petit voisinage compact  $U$  de chaque point fixe, et qui n'a dans  $U$  qu'un seul point fixe, qui est hyperbolique. Par partition de l'unité et carte locale, il suffit de travailler avec  $M = \mathbb{R}^N$  et  $\phi$  difféomorphisme local en 0, fixant 0, et de différentielle en 0 n'ayant pas la valeur propre 1. Pour construire une application  $\varphi_\epsilon$  comme dans l'énoncé, il suffit de multiplier l'application identité par une fonction  $x \mapsto f_\epsilon(\|x\|)$  où  $f_\epsilon : [0, +\infty[ \rightarrow \mathbb{R}$  est  $C^\infty$ , croissante, vaut 1 sur  $]2\epsilon, +\infty[$  et  $1 - \epsilon$  sur  $[0, \epsilon]$ , et dont les dérivées convergent uniformément vers 0 quand  $\epsilon \rightarrow 0$ .

Notons que 0 est un point fixe dans  $U$  de  $\phi \circ \varphi_\epsilon$ , donc c'est le seul par l'ouverture montrée à la question (3) si  $\epsilon$  est assez petit. Comme  $d_0(\phi \circ \varphi_\epsilon) = (1 - \epsilon)d_0\phi$ , il suffit donc de prendre  $\epsilon$  arbitrairement petit tel que  $d_0\phi$  n'ait pas de valeur propre de module  $\frac{1}{1-\epsilon}$ .

(5) Pour tout  $n \in \mathbb{N} - \{0\}$ , notons  $\mathcal{A}_n$  l'ensemble des  $C^k$ -difféomorphismes  $\phi$  de  $M$  ayant un nombre fini de points périodiques de période  $n$ , en lesquels la différentielle de  $\phi^n$  soit hyperbolique. Montrons que  $\mathcal{A}_n$  est un ouvert dense de  $\text{Diff}^k(M)$ . Comme  $\mathcal{A}_\infty = \bigcap_{n \in \mathbb{N} - \{0\}} \mathcal{A}_n$  est un  $G_\delta$  dense puisque  $\text{Diff}^k(M)$  est un espace de Baire, ceci montre la question (5). Le cas  $n = 1$  a été traité dans la question (4). Le caractère ouvert de  $\mathcal{A}_n$  découle du cas  $n = 1$ , car si  $\phi'$  est suffisamment  $C^k$ -proche de  $\phi$ , alors  $\phi'^n$  est  $C^k$ -proche de  $\phi^n$ . Nous renvoyons à [KH] pour la densité.

(6) Soit  $\phi' \in \mathcal{A}_\infty$  arbitrairement proche de  $\phi$ . Puisque  $\phi$  est structurellement stable, les difféomorphismes  $\phi'$  et  $\phi$  sont conjugués. Puisque qu'une conjugaison induit une bijection de l'ensemble des points fixes de période  $n$  de  $\phi$  sur celui de  $\phi'$ , le résultat en découle.

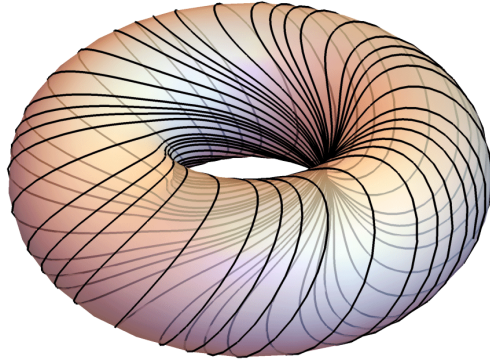
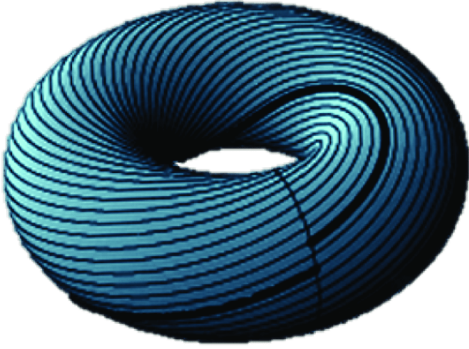
**Correction de l'exercice E.62.** (1) Il est immédiat que 0 est un point fixe de  $\phi$  car  $\phi(0) = \mathcal{M}0 \pmod{\mathbb{Z}^2} = 0$ .

Notons  $\phi = \frac{1+\sqrt{5}}{2}$  le nombre d'or et  $\bar{\phi} = \frac{1-\sqrt{5}}{2}$  son conjugué de Galois. L'espace tangent  $T_0\mathbb{T}^2$  en 0 (et en fait en tout point du tore  $\mathbb{T}^2$ ) s'identifie à  $\mathbb{R}^2$ . Puisque la différentielle en tout point d'une application linéaire est égale à elle-même, l'application tangente  $T_0\phi : T_0\mathbb{T}^2 \rightarrow T_0\mathbb{T}^2$  s'identifie donc à l'application linéaire  $\mathcal{M} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Les valeurs propres de  $\mathcal{M}$  sont  $\phi^2 = \frac{3+\sqrt{5}}{2} > 1$  et  $\bar{\phi}^2 = \frac{3-\sqrt{5}}{2} \in ]0, 1[$ , correspondant aux espaces propres  $E_0^u = \mathbb{R}(1, \phi^2 - 2)$  (sur lequel  $T_0\phi$  agit par dilatation par  $\phi^2$ ) et  $E_0^s = \mathbb{R}(1, \bar{\phi}^2 - 2)$  (sur lequel  $T_0\phi$  agit par dilatation par  $\bar{\phi}^2$ ).

(2) Puisque la projection canonique  $\gamma : \mathbb{R}^2 \rightarrow \mathbb{T}^2$  est une isométrie locale, et puisque les variétés stables/instables locales de  $\phi$  sont des parties où  $\phi$  contracte/dilata (voir les formules (3)), les variétés stables et instables locales de  $\phi$  en 0 sont des images par  $\gamma$  de segments ouverts contenant 0 dans  $E_0^s$  et  $E_0^u$  respectivement.

Par la définition des variétés stables/instables globales en fonction de celles locales (voir les formules (126)), nous avons donc que  $W^s(0) = \gamma(E_0^s)$  et  $W^u(0) = \gamma(E_0^u)$ . Ces variétés stables et instables sont des droites immergées denses dans  $\mathbb{T}^2$ , car  $E_0^s$  et  $E_0^u$  sont des droites de pente irrationnelle.





## 11 Mesures d'entropie maximale

L'entropie topologique est une mesure de la complexité d'un système dynamique topologique, qui n'est pertinente que lorsqu'il est "chaotique". Si nous comptons le nombre minimum d'orbites de longueur  $n$  qui permet de les "pister" toutes à  $\epsilon$ -près, nous obtenons une fonction de  $n$ . L'entropie topologique est, grosso modo, le taux de croissance exponentiel de cette fonction de  $n$  lorsque  $\epsilon$  est très petit.

Le premier but de ce chapitre est de définir précisément l'entropie topologique (partie 11.1) et de montrer (partie 11.2) qu'elle est égale à la borne supérieure des entropies mesurées pour toutes les mesures de probabilités boréliennes invariantes, lorsque l'espace des phases est compact. Nous renvoyons à la partie 5 ou [PY, Wal, CorFS, Pet, EL, Cou] pour tout ce qui concerne l'entropie mesurée des systèmes dynamiques mesurés. Nous nous intéresserons au problème de l'existence et de l'unicité des mesures d'entropie maximale : ces mesures apportent des informations fondamentales sur la structure statistique du système dynamique considéré.

Nous calculerons l'entropie topologique de quelques exemples dans les deux parties 11.3 pour les systèmes dynamiques symboliques et 11.4 pour un exemple représentatif de systèmes dynamiques hyperboliques, par la méthode du codage par partition de Markov. Nous montrerons en particulier que les mesures d'entropie maximale jouent un rôle crucial afin de comprendre la distribution des points périodiques des systèmes dynamiques hyperboliques. Les problèmes d'équidistribution de points périodiques ont donné lieu à de très nombreux développements (voir par exemple le survol de Sharp dans [Mar3], [DiS, AkS, Tak] et [BrPP, Part II]).

Des références possibles pour ce chapitre sont les livres [Wal, KH, PY].

### 11.1 Entropie topologique

Une manière possible de mesurer la complexité d'un espace topologique est de comprendre comment se comportent ses recouvrements ouverts de plus en plus "fins". Le comportement de ces recouvrements ouverts fins par itération d'une application continue permettra de mesurer le "désordre" d'un système dynamique topologique par un invariant numérique, que nous appellerons l'entropie topologique.

#### Recouvrements ouverts

Soit  $X$  un espace topologique compact. Pour tout recouvrement ouvert  $\mathcal{U}$  de  $X$ , notons  $N(\mathcal{U})$  le cardinal minimum des sous-recouvrements finis de  $\mathcal{U}$ . (L'existence de cet entier  $N(\mathcal{U})$  résulte de la compacité de  $X$ .) Pour toute application continue  $\phi : X' \rightarrow X$  où  $X'$  est un espace topologique compact, notons  $\phi^{-1}\mathcal{U}$  le recouvrement ouvert  $\{\phi^{-1}(U) : U \in \mathcal{U}\}$  de  $X'$ .

Soit  $\mathcal{V}$  un (autre) recouvrement ouvert de  $X$ . Notons  $\mathcal{U} \vee \mathcal{V}$  le recouvrement ouvert de  $X$ , appelé le *joint* de  $\mathcal{U}$  et  $\mathcal{V}$ , défini par

$$\mathcal{U} \vee \mathcal{V} = \{U \cap V : U \in \mathcal{U}, V \in \mathcal{V}\}.$$

L'opérateur binaire  $\vee$  est une loi de composition interne sur l'ensemble  $\mathcal{R}$  des recouvrements

ouverts de  $X$ , qui est associative<sup>239</sup> et commutative<sup>240</sup>. Pour  $\mathcal{U}_1, \dots, \mathcal{U}_n \in \mathcal{R}$ , notons

$$\bigvee_{i=1}^n \mathcal{U}_i = \mathcal{U}_1 \vee \dots \vee \mathcal{U}_n = \{U_1 \cap \dots \cap U_n : \forall i = 1, \dots, n, U_i \in \mathcal{U}_i\}.$$

Nous dirons que  $\mathcal{U}$  est *moins fin* que  $\mathcal{V}$  (ou que  $\mathcal{V}$  est *plus fin que*  $\mathcal{U}$ ), et nous noterons  $\mathcal{U} \preceq \mathcal{V}$ , si tout ouvert de  $\mathcal{V}$  est contenu dans un ouvert de  $\mathcal{U}$ . La relation  $\preceq$  est une relation d'ordre sur l'ensemble des recouvrements ouverts de  $X$ . Notons que  $\mathcal{U} \preceq \mathcal{U} \vee \mathcal{U}$ , mais qu'en général, nous avons  $\mathcal{U} \vee \mathcal{U} \neq \mathcal{U}$ .

Voici quelques propriétés élémentaires de cet entier  $N(\mathcal{U})$ .

**Lemme 11.1.** *Les propriétés suivantes sont satisfaites.*

- i) Nous avons  $N(\phi^{-1}\mathcal{U}) \leq N(\mathcal{U})$  avec égalité lorsque  $\phi$  est surjective.
- ii) Nous avons  $\phi^{-1}(\mathcal{U} \vee \mathcal{V}) = (\phi^{-1}\mathcal{U}) \vee (\phi^{-1}\mathcal{V})$  et  $N(\mathcal{U} \vee \mathcal{V}) \leq N(\mathcal{U}) N(\mathcal{V})$ .
- iii) Si  $\mathcal{U} \preceq \mathcal{V}$ , alors  $\phi^{-1}\mathcal{U} \preceq \phi^{-1}\mathcal{V}$  et  $N(\mathcal{U}) \leq N(\mathcal{V})$ .

**Démonstration.** i) Si les ouverts  $U_1, \dots, U_n$  appartenant à  $\mathcal{U}$  recouvrent l'espace  $X$ , alors les ouverts  $\phi^{-1}(U_1), \dots, \phi^{-1}(U_n)$  appartenant à  $\phi^{-1}\mathcal{U}$  recouvrent  $X'$ . La réciproque est vraie lorsque  $\phi$  est surjective.

ii) Si les ouverts  $U_1, \dots, U_n$  appartenant à  $\mathcal{U}$  recouvrent  $X$  et si les ouverts  $V_1, \dots, V_p$  appartenant à  $\mathcal{V}$  recouvrent  $X$ , alors les ouverts  $U_i \cap V_j$  pour  $i = 1, \dots, n$  et  $j = 1, \dots, p$  recouvrent aussi  $X$ .

iii) Si les ouverts  $V_1, \dots, V_p$  appartenant à  $\mathcal{V}$  recouvrent  $X$ , choisissons pour tout  $i$  un ouvert  $U_i$  de  $\mathcal{U}$  contenant  $V_i$ . Alors les ouverts  $U_1, \dots, U_p$  de  $\mathcal{U}$  recouvrent aussi  $X$ .  $\square$

Soit maintenant  $(X, d)$  un espace métrique compact. Les exemples les plus simples de recouvrements de  $X$  sont les recouvrements  $\mathcal{O}_\epsilon$  par les boules ouvertes de rayon  $\epsilon$  dans  $X$ , où  $\epsilon > 0$ , définis par

$$\mathcal{O}_\epsilon = \{ B(x, \epsilon) = \{y \in X : d(x, y) < \epsilon\} : x \in X \}.$$

Nous dirons qu'une partie  $F$  de  $X$  est  $\epsilon$ -dense (pour  $d$ ) si tout point de  $X$  est à distance strictement inférieure à  $\epsilon$  d'au moins un point de  $F$ , c'est-à-dire si nous avons

$$X = \bigcup_{x \in F} B(x, \epsilon).$$

Nous dirons que  $F$  est  $\epsilon$ -séparée (pour  $d$ ) si deux points distincts de  $F$  sont à distance au moins  $\epsilon$  l'un de l'autre, c'est-à-dire si nous avons

$$\inf \{d(x, y) : x, y \in F, x \neq y\} \geq \epsilon.$$

Notons  $N_d(\epsilon)$  la borne inférieure des cardinaux des parties  $\epsilon$ -denses de  $X$ , et  $N'_d(\epsilon)$  la borne supérieure des cardinaux des parties  $\epsilon$ -séparées de  $X$ . Ces deux nombres sont finis, par compacité de  $X$ . Nous avons bien sûr  $N_d(\epsilon) = N(\mathcal{O}_\epsilon)$ .

**Lemme 11.2.** *Pour tout  $\epsilon > 0$ , nous avons les inégalités*

$$N'_d(2\epsilon) \leq N_d(\epsilon) \leq N'_d(\epsilon).$$

239. pour tous les  $\mathcal{U}, \mathcal{V}, \mathcal{W} \in \mathcal{R}$ , nous avons  $(\mathcal{U} \vee \mathcal{V}) \vee \mathcal{W} = \mathcal{U} \vee (\mathcal{V} \vee \mathcal{W})$

240. pour tous les  $\mathcal{U}, \mathcal{V} \in \mathcal{R}$ , nous avons  $\mathcal{U} \vee \mathcal{V} = \mathcal{V} \vee \mathcal{U}$

**Démonstration.** En effet, d'une part deux points d'une partie  $(2\epsilon)$ -séparée n'appartiennent pas à une même boule ouverte de rayon  $\epsilon$  (par l'inégalité triangulaire). D'autre part, une partie  $\epsilon$ -séparée de cardinal maximal est forcément  $\epsilon$ -dense, sinon nous pourrions lui rajouter un point en la laissant  $\epsilon$ -séparée.  $\square$

### Définition de l'entropie topologique

Soit  $(X, \phi)$  un système dynamique topologique à temps discret, où  $X$  est compact.

**Lemme 11.3.** *Pour tout recouvrement ouvert  $\mathcal{U}$  de  $X$ , la suite  $(\frac{1}{n} \ln N(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{U}))_{n \in \mathbb{N}}$  admet une limite finie quand  $n$  tends vers  $+\infty$ .*

**Démonstration.** La suite  $(a_n = \ln N(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{U}))_{n \in \mathbb{N}}$  est sous-additive, car par le lemme 11.1 ii) et i), nous avons

$$\begin{aligned} a_{n+m} &\leq \log N\left(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{U}\right) + \log N\left(\bigvee_{i=n}^{n+m-1} \phi^{-i}\mathcal{U}\right) \\ &= \log N\left(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{U}\right) + \log N\left(\phi^{-n}\left(\bigvee_{i=0}^{m-1} \phi^{-i}\mathcal{U}\right)\right) \leq a_n + a_m. \end{aligned}$$

Le résultat découle alors du lemme classique dit de Fekete 2.7 (voir aussi le lemme 8.2) sur les suites sous-additives.  $\square$

La limite donnée par le lemme 11.3 est appelée l'entropie de  $\phi$  relativement au recouvrement ouvert  $\mathcal{U}$ , et nous la noterons

$$h(\phi, \mathcal{U}) = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln N\left(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{U}\right) = \inf_{n \in \mathbb{N} - \{0\}} \frac{1}{n} \ln N\left(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{U}\right).$$

Si  $\phi$  est inversible, par le lemme 11.1 i), nous avons donc

$$h(\phi, \mathcal{U}) = \lim_{n \rightarrow +\infty} \frac{1}{2n+1} \ln N\left(\bigvee_{i=-n}^n \phi^{-i}\mathcal{U}\right) = \inf_{n \in \mathbb{N} - \{0\}} \frac{1}{2n+1} \ln N\left(\bigvee_{i=-n}^n \phi^{-i}\mathcal{U}\right). \quad (133)$$

**Remarque.** Si  $\mathcal{U} \preceq \mathcal{V}$ , alors  $h(\phi, \mathcal{U}) \leq h(\phi, \mathcal{V})$ , donc passer à un recouvrement plus fin ne peut que faire croître l'entropie relativement au recouvrement. Ceci motive la définition suivante.

Nous appellerons *entropie topologique* de  $\phi$ , et nous noterons  $h_{\text{top}}(\phi)$  l'élément de  $[0, +\infty]$  suivant :

$$h_{\text{top}}(\phi) = \sup_{\mathcal{U}} h(\phi, \mathcal{U}),$$

où la borne supérieure est prise sur tous les recouvrements ouverts  $\mathcal{U}$  de  $X$ .

La proposition suivante dit que l'entropie topologique est un invariant de conjugaison des systèmes dynamiques topologiques.

**Proposition 11.4.** *Soient  $(X, \phi)$  et  $(X', \phi')$  des systèmes dynamiques topologiques à temps discrets (topologiquement) conjugués, avec  $X$  et  $X'$  compacts. Alors*

$$h_{\text{top}}(\phi) = h_{\text{top}}(\phi').$$

**Démonstration.** Soit  $\mathfrak{h} : X' \rightarrow X$  un homéomorphisme tel que  $\mathfrak{h} \circ \phi' = \phi \circ \mathfrak{h}$ . Pour tout recouvrement ouvert  $\mathcal{V}$  de  $X$ , le recouvrement ouvert  $\mathfrak{h}^{-1}\mathcal{V} = \{\mathfrak{h}^{-1}(V) : V \in \mathcal{V}\}$  de  $X'$  vérifie  $N(\mathfrak{h}^{-1}\mathcal{V}) = N(\mathcal{V})$ . Donc  $h(\phi, \mathcal{V}) = h(\phi', \mathfrak{h}^{-1}\mathcal{V})$ . D'où  $h_{\text{top}}(\phi) \leq h_{\text{top}}(\phi')$ , et par symétrie, le résultat en découle.  $\square$

Si  $X$  est métrisable, fixons une distance  $d$  sur  $X$  compatible avec la topologie de  $X$ . La proposition suivante fournit une définition plus intuitive de l'entropie topologique de  $\phi$  : elle affirme que nous pouvons nous contenter des recouvrements  $\mathcal{O}_\epsilon$ .

Notons  $d_n^\phi$  la distance sur  $X$ , appelée la *distance de pistage* jusqu'au temps  $n$ , définie par

$$d_n^\phi(x, y) = \max_{0 \leq i \leq n-1} d(\phi^i(x), \phi^i(y)).$$

L'entier  $N_{d_n^\phi}(\epsilon)$  peut être considéré comme le nombre minimum d'orbites de longueur  $n$  permettant de pister à  $\epsilon$  près n'importe quelle orbite de longueur  $n$ . L'entier  $N'_{d_n^\phi}(\epsilon)$  est à considérer comme le nombre maximum d'orbites de longueur  $n$  telles que deux d'entre elles ne se suivent pas à  $\epsilon$  près. Nous avons l'égalité

$$N_{d_n^\phi}(\epsilon) = N\left(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{O}_\epsilon\right).$$

**Proposition 11.5.** *Pour toute distance  $d$  sur  $X$  induisant sa topologie, nous avons*

$$h_{\text{top}}(\phi) = \lim_{\epsilon \rightarrow 0} \left( \lim_{n \rightarrow \infty} \frac{1}{n} \ln N_{d_n^\phi}(\epsilon) \right) = \lim_{\epsilon \rightarrow 0} \left( \limsup_{n \rightarrow \infty} \frac{1}{n} \ln N'_{d_n^\phi}(\epsilon) \right).$$

**Remarque.** En particulier, puisque  $h_{\text{top}}(\phi)$  est indépendante de  $d$ , ces limites restent inchangées si nous remplaçons la distance  $d$  par une autre distance qui induit la même topologie.

**Démonstration.** La première égalité résulte du lemme suivant, et la seconde du lemme 11.2.  $\square$

**Lemme 11.6. (Lemme de Lebesgue)** *Pour tout recouvrement ouvert  $\mathcal{U}$  d'un espace métrique compact  $X$ , il existe  $\epsilon > 0$  tel que  $\mathcal{U}$  est moins fin que  $\mathcal{O}_\epsilon$ .*

**Démonstration.** Pour tout  $x$  dans  $X$ , il existe  $\epsilon_x > 0$  tel que la boule ouverte  $B(x, 2\epsilon_x)$  soit contenue dans l'un des ouverts de  $\mathcal{U}$ . Nous extrayons du recouvrement de  $X$  par ses boules ouvertes  $B(x, \epsilon_x)$  un sous-recouvrement fini  $(B(x, \epsilon_x))_{x \in F}$ , et nous prenons  $\epsilon = \inf_{x \in F} \epsilon_x > 0$ .  $\square$

Si  $\mathcal{U}$  est un recouvrement ouvert de  $X$ , notons  $\text{diam}(\mathcal{U})$  la borne supérieure des diamètres (pour la distance  $d$  fixée ci-dessus) des ouverts  $U$  de  $\mathcal{U}$ .

**Corollaire 11.7.** *Si  $X$  est métrisable, pour toute suite  $(\mathcal{U}_n)_{n \in \mathbb{N}}$  de recouvrements ouverts de  $X$  telle que  $\lim_{n \rightarrow \infty} \text{diam}(\mathcal{U}_n) = 0$ , nous avons*

$$h_{\text{top}}(\phi) = \lim_{n \rightarrow \infty} h(\phi, \mathcal{U}_n).$$

**Démonstration.** En effet, pour tout  $\epsilon > 0$ , le recouvrement ouvert  $\mathcal{U}_n$  est plus fin que le recouvrement ouvert  $\mathcal{O}_\epsilon$  dès que  $\text{diam}(\mathcal{U}_n) < \epsilon$ , donc  $h(\phi, \mathcal{U}_n) \geq h(\phi, \mathcal{O}_\epsilon)$ .  $\square$

### Propriétés élémentaires de l'entropie topologique

Soit  $(X, \phi)$  un système dynamique topologique à temps discret, où  $X$  est compact.

**Proposition 11.8.** Soit  $\mathcal{U}$  un recouvrement ouvert de  $X$ .

- (1) Pour tout  $k \geq 1$ , nous avons  $h(\phi, \mathcal{U}) = h(\phi, \bigvee_{i=0}^k \phi^{-i} \mathcal{U})$ , et, si  $\phi$  est inversible, alors  $h(\phi, \mathcal{U}) = h(\phi, \bigvee_{i=-k}^k \phi^{-i} \mathcal{U})$ .
- (2) Pour tout  $k \in \mathbb{N}$ , nous avons  $h_{\text{top}}(\phi^k) = k h_{\text{top}}(\phi)$ . Si  $\phi$  est inversible, alors nous avons  $h_{\text{top}}(\phi^k) = |k| h_{\text{top}}(\phi)$  pour tout  $k \in \mathbb{Z}$ .

**Démonstration.** (1) Pour tout  $n \in \mathbb{N}$ , notons  $a_n(\mathcal{U}) = \log(N(\bigvee_{i=0}^n \phi^{-i} \mathcal{U}))$ . Nous avons  $a_n(\bigvee_{i=0}^k \phi^{-i} \mathcal{U}) = a_{n+k}(\mathcal{U})$ , et lorsque  $\phi$  est inversible,  $a_n(\bigvee_{i=-k}^k \phi^{-i} \mathcal{U}) = a_{n+2k-1}(\mathcal{U})$ . Le résultat s'en déduit, car pour tout  $k' \in \mathbb{N}$  les suites  $(\frac{a_n(\mathcal{U})}{n})_{n \in \mathbb{N}}$  et  $(\frac{a_{n+k'}(\mathcal{U})}{n})_{n \in \mathbb{N}}$  ont la même limite quand  $n$  tend vers  $+\infty$ .

(2) Si  $k = 0$ , alors la suite  $(N(\bigvee_{i=0}^{n-1} \mathcal{U}))_{n \in \mathbb{N}}$  est bornée, car le recouvrement ouvert  $\bigvee_{i=0}^{n-1} \mathcal{U}$  contient le recouvrement ouvert  $\mathcal{U}$ , donc

$$h(\text{id}, \mathcal{U}) = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln N\left(\bigvee_{i=0}^{n-1} \mathcal{U}\right) = 0$$

pour tout recouvrement ouvert  $\mathcal{U}$ . Si  $k \geq 1$ , nous avons

$$h(\phi^k, \mathcal{U}) \leq h(\phi^k, \bigvee_{i=0}^k \phi^{-i} \mathcal{U}) = \lim_{n \rightarrow \infty} \frac{1}{n} a_{nk}(\mathcal{U}) = k h(\phi, \mathcal{U}) \leq k h_{\text{top}}(\phi).$$

Les inégalités montrent que  $h(\phi^k, \mathcal{U}) \leq k h_{\text{top}}(\phi)$ , et les égalités (prises en sens inverse!) que  $k h_{\text{top}}(\phi) \leq h_{\text{top}}(\phi^k)$ . Donc  $h_{\text{top}}(\phi^k) = k h_{\text{top}}(\phi)$ . La dernière affirmation découle de l'exercice suivant.  $\square$

**Exercice E.64.** Soit  $\phi$  un homéomorphisme d'un espace métrique compact  $X$ . Montrer que  $h_{\text{top}}(\phi^{-1}) = h_{\text{top}}(\phi)$ .

**Remarque.** La proposition 11.8 (2) rend naturelle la définition suivante. Soit  $\psi = (\psi^t)_{t \in \mathbb{R}}$  un groupe à un paramètre continu<sup>241</sup> d'homéomorphismes d'un espace topologique compact  $X$ . L'entropie topologique du système dynamique topologique à temps continu  $(X, \psi)$  est définie par

$$h_{\text{top}}(\psi) = h_{\text{top}}(\psi^1).$$

**Exercice E.65.** Montrer que  $h_{\text{top}}(\psi^t) = |t| h_{\text{top}}(\psi)$  pour tout  $t \in \mathbb{R}$ .

Supposons  $X$  muni d'une distance  $d$  induisant sa topologie. Le résultat suivant permet d'éviter de prendre la borne supérieure dans la définition de l'entropie topologique.

Un recouvrement ouvert  $\mathcal{U}$  de  $X$  est dit *générateur* pour  $\phi$  si

$$\lim_{n \rightarrow \infty} \text{diam}\left(\bigvee_{i=0}^n \phi^{-i} \mathcal{U}\right) = 0.$$

Lorsque  $\phi$  est inversible, nous demanderons seulement que

$$\lim_{n \rightarrow \infty} \text{diam}\left(\bigvee_{i=-n}^n \phi^{-i} \mathcal{U}\right) = 0.$$

241. Rappelons que ceci signifie que l'application de  $\mathbb{R} \times X$  dans  $X$  définie par  $(t, x) \mapsto \psi^t(x)$  est continue.

**Théorème 11.9.** *Si un recouvrement ouvert  $\mathcal{U}$  est générateur pour  $\phi$ , alors*

$$h_{\text{top}}(\phi) = h(\phi, \mathcal{U}) .$$

**Démonstration.** D'après la proposition 11.8 (1), la suite  $(h(\phi, \bigvee_{i=0}^n \phi^{-i}\mathcal{U}))_{n \in \mathbb{N}}$  est constante égale à  $h(\phi, \mathcal{U})$ . D'après le corollaire 11.7, cette suite converge vers  $h_{\text{top}}(\phi)$ . Nous avons donc  $h_{\text{top}}(\phi) = h(\phi, \mathcal{U})$ . Nous procédons de même lorsque  $\phi$  est inversible, avec la suite  $(h(\phi, \bigvee_{i=-n}^n \phi^{-i}\mathcal{U}))_{n \in \mathbb{N}}$ .  $\square$

**Remarque.** Nous renvoyons aux lignes précédant la proposition 10.19 pour la définition d'un système dynamique topologique expansif. Notons que le système dynamique  $(X, \phi)$  est expansif si et seulement s'il existe  $\epsilon_0 > 0$  (une constante d'expansivité de  $(X, \phi)$ ) tel que le recouvrement ouvert  $\mathcal{O}_{\epsilon_0}$  de  $X$  soit générateur pour  $\phi$ .

**Corollaire 11.10.** *L'entropie topologique  $h_{\text{top}}(\phi)$  d'un système dynamique topologique expansif  $(X, \phi)$ , où  $X$  est un espace métrique compact, est finie : nous avons*

$$h_{\text{top}}(\phi) < \infty .$$

De plus, si  $\epsilon_0$  est une constante d'expansivité de  $\phi$ , alors, pour tout  $\epsilon \in ]0, \epsilon_0[$ , nous avons

$$h_{\text{top}}(\phi) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln N_{d_n^\phi}(\epsilon) = \limsup_{n \rightarrow \infty} \frac{1}{n} \ln N'_{d_n^\phi}(\epsilon) .$$

**Démonstration.** En effet,  $h_{\text{top}}(\phi) = h(\phi, \mathcal{O}_{\epsilon_0})$  où  $\epsilon_0 > 0$  est tel que le recouvrement ouvert  $\mathcal{O}_{\epsilon_0}$  soit générateur. Nous utilisons alors la proposition 11.8 (1) et l'égalité précédant la proposition 11.5.  $\square$

## 11.2 Principe variationnel

Soit  $(X, \phi)$  un système dynamique topologique, où  $X$  est un espace métrique compact. Rappelons (voir la partie 0.3) que nous notons  $\text{Prob}_\phi(X)$  l'espace métrisable compact des mesures de probabilité boréliennes sur  $X$  invariantes par  $\phi$ , muni de la topologie vague (égale à la topologie faible-étoile).

**Théorème 11.11. (Principe variationnel)** *Soient  $X$  un espace métrique compact et  $\phi : X \rightarrow X$  une application continue. Alors nous avons l'égalité*

$$h_{\text{top}}(\phi) = \sup_{\mu \in \text{Prob}_\phi(X)} h_\mu(\phi) .$$

**Démonstration.** Nous procédons en deux étapes.

**Étape 1.** Montrons la majoration  $h_\mu(\phi) \leq h_{\text{top}}(\phi)$  pour tout  $\mu$  dans  $\text{Prob}_\phi(X)$ .

Nous aurons besoin du lemme suivant, qui est un des lemmes clefs de la démonstration du théorème de représentation de Riesz, voir par exemple [Coh, Prop. 7.2.3].

**Lemme 11.12.** *Soient  $X$  un espace métrisable localement compact séparable,  $\mu$  une mesure de probabilité borélienne sur  $X$  et  $A$  un borélien de  $X$ . Alors, pour tout  $\epsilon > 0$ , il existe un compact  $K$  de  $X$  et un ouvert  $U$  de  $X$  tels que  $K \subset A \subset U$  et  $\mu(U - K) \leq \epsilon$ .  $\square$*

Soit  $\alpha = \{A_1, \dots, A_p\}$  une partition mesurable de  $X$ . D'après le lemme ci-dessus, il existe des compacts  $K_1, \dots, K_p$  contenus dans  $A_1, \dots, A_p$  respectivement (donc deux à deux disjoints) tels que le nombre réel  $\epsilon = \sup_{1 \leq i \leq n} \mu(A_i - K_i) \geq 0$  soit aussi petit que nous le souhaitons. Posons  $U_0 = X - (K_1 \cup \dots \cup K_p)$ , et introduisons la partition mesurable  $\beta = \{K_1, \dots, K_p, U_0\}$ . D'après la proposition 5.1 (7), nous pouvons nous arranger pour que  $H_\mu(\alpha|\beta) \leq 1$ . L'inégalité de Rokhlin (voir la proposition 5.3 (4)) assure que

$$h_\mu(\phi, \alpha) \leq h_\mu(\phi, \beta) + 1.$$

Remarquons que  $U_0 \cup K_i = X - (K_1 \cup \dots \cup \widehat{K_i} \cup \dots \cup K_p)$  est un ouvert de  $X$  pour tout  $i = 1, \dots, p$ . Introduisons maintenant le recouvrement ouvert  $\mathcal{U} = \{U_0 \cup K_1, \dots, U_0 \cup K_p\}$  de  $X$ .

**Lemme 11.13.** *Pour tout  $n \in \mathbb{N} - \{0\}$ , nous avons  $\text{Card} \bigvee_{i=0}^{n-1} \phi^{-i} \beta \leq 2^n N \left( \bigvee_{i=0}^{n-1} \phi^{-i} \mathcal{U} \right)$ .*

**Démonstration.** Pour tous les  $i_0, \dots, i_{n-1} \in \{0, \dots, p\}$ , notons  $K_0 = U_0$  et

$$K_{i_0, \dots, i_{n-1}} = K_{i_0} \cap \phi^{-1}(K_{i_1}) \cap \dots \cap \phi^{-(n-1)}(K_{i_{n-1}}).$$

Alors les  $K_{i_0, \dots, i_{n-1}}$  sont deux à deux disjoints et

$$\text{Card} \left( \bigvee_{i=0}^{n-1} \phi^{-i} \beta \right) = \text{Card} \{ (i_0, \dots, i_{n-1}) \in \{0, \dots, p\}^n : K_{i_0, \dots, i_{n-1}} \neq \emptyset \}.$$

Le lemme découle alors du fait que  $K_{i_0, \dots, i_{n-1}}$  est contenu dans

$$(U_0 \cup K_{\max\{i_0, 1\}}) \cap \phi^{-1}(U_0 \cup K_{\max\{i_1, 1\}}) \cap \dots \cap \phi^{-(n-1)}(U_0 \cup K_{\max\{i_{n-1}, 1\}}),$$

ce qui donne  $2^n$  choix possibles d'un élément du recouvrement ouvert  $\bigvee_{i=0}^{n-1} \phi^{-i} \mathcal{U}$  pour tout évènement de la partition mesurable  $\bigvee_{i=0}^{n-1} \phi^{-i} \beta$  (ces évènements étant deux à deux disjoints).  $\square$

D'après la proposition 5.1 (1) et en passant au logarithme le lemme ci-dessus, nous avons donc la majoration

$$H_\mu \left( \bigvee_{i=0}^{n-1} \phi^{-i} \beta \right) \leq \ln \left( \text{Card} \left( \bigvee_{i=0}^{n-1} \phi^{-i} \beta \right) \right) \leq \ln N \left( \bigvee_{i=0}^{n-1} \phi^{-i} \mathcal{U} \right) + n \ln 2.$$

En divisant par  $n$  et en passant à la limite, nous obtenons

$$h_\mu(\phi, \alpha) \leq h_\mu(\phi, \beta) + 1 \leq h(\phi, \mathcal{U}) + \log 2 + 1 \leq h_{\text{top}}(\phi) + \log 2 + 1.$$

Nous en déduisons la majoration  $h_\mu(\phi) \leq h_{\text{top}}(\phi) + \log 2 + 1$ . En remplaçant dans cette égalité  $\phi$  par une puissance  $\phi^n$  où  $n \in \mathbb{N} - \{0\}$ , en utilisant les propositions 5.3 (6) et 11.8 (2), et en divisant par  $n$ , nous obtenons  $h_\mu(\phi) \leq h_{\text{top}}(\phi) + \frac{\log 2 + 1}{n}$ . Par conséquent, nous avons  $h_\mu(\phi) \leq h_{\text{top}}(\phi)$ , ce qui conclut l'étape 1.

**Étape 2.** Montrons que  $\sup_{\mu \in \text{Prob}_\phi(X)} h_\mu(\phi) \geq h_{\text{top}}(\phi)$ , ce qui avec l'étape 1 démontre le principe variationnel.

Le lemme suivant est l'un des sens du résultat classique (voir le théorème de Portmanteau dans [Bil]) qui dit que la convergence en loi (qui est la même que la convergence étroite et que la convergence vague lorsque l'espace est métrisable compact) des mesures de probabilité est équivalente à la convergence des mesures des boréliens donnés de frontière de mesure nulle.



**Lemme 11.14.** Soient  $X$  un espace métrisable compact,  $(\mu_n)_{n \in \mathbb{N}}$  une suite de mesures de probabilité boréliennes sur  $X$  qui convergent vaguement vers une mesure de probabilité  $\mu$  et  $A$  un borélien de  $X$ , tel que  $\mu(\partial A) = 0$ . Alors

$$\lim_{n \rightarrow \infty} \mu_n(A) = \mu(A) .$$

**Démonstration.** Comme l'adhérence  $\bar{A}$  de  $A$  est fermée, la fonction caractéristique  $\mathbb{1}_{\bar{A}}$  est une limite d'une suite décroissante de fonctions continues  $(f_k)_{k \in \mathbb{N}}$  : nous pouvons par exemple prendre  $f_k(x) = \max\{1 - k d(x, \bar{A}), 0\}$ . L'inégalité  $\mu_n(\bar{A}) \leq \mu_n(f_k)$  et les égalités  $\lim_{n \rightarrow \infty} \mu_n(f_k) = \mu(f_k)$  et  $\lim_{k \rightarrow \infty} \mu(f_k) = \mu(\bar{A})$  (cette dernière par le théorème de convergence monotone de Lebesgue) impliquent que

$$\limsup_{n \rightarrow \infty} \mu_n(\bar{A}) \leq \mu(\bar{A}) .$$

De même, l'intérieur  $\overset{\circ}{A}$  de  $A$  vérifie  $\liminf_{n \rightarrow \infty} \mu_n(\overset{\circ}{A}) \geq \mu(\overset{\circ}{A})$ . L'hypothèse du lemme assure que  $\mu(\bar{A}) = \mu(\overset{\circ}{A}) = \mu(A)$ . Donc  $\lim_{n \rightarrow \infty} \mu_n(A) = \mu(A)$ .  $\square$

**Lemme 11.15.** Soient  $X$  un espace métrisable compact,  $\mu$  une mesure de probabilité borélienne sur  $X$  et  $\epsilon > 0$ . Alors il existe une partition mesurable  $\alpha$  de  $X$  telle que, pour tout  $A$  dans  $\alpha$ , nous avons  $\text{diam}(A) \leq \epsilon$  et  $\mu(\partial A) = 0$ .

**Démonstration.** Par la finitude de  $\mu$ , pour tout point  $x$  de  $X$ , il existe un réel  $\epsilon_x$  dans  $]0, \frac{\epsilon}{2}[$  tel que la sphère de centre  $x$  et de rayon  $\epsilon_x$  dans  $X$  soit de mesure nulle pour  $\mu$ . Nous pouvons extraire du recouvrement ouvert de  $X$  par les boules  $B(x, \epsilon_x)$  pour  $x \in X$  un sous-recouvrement fini  $\{B_1, \dots, B_p\}$ . Par construction, nous avons  $\mu(\partial B_i) = 0$ . Nous construisons  $\alpha = \{A_1, \dots, A_p\}$  en posant  $A_i = B_i - (B_1 \cup \dots \cup B_{i-1})$ . Nous avons alors  $\partial A_i \subset \partial B_1 \cup \dots \cup \partial B_p$ , donc  $\mu(\partial A_i) = 0$ .  $\square$

Le lemme crucial suivant permet de construire des mesures de probabilités boréliennes invariantes par  $\phi$  de grande entropie mesurée. Notons  $\Delta_x$  la masse de Dirac unité en tout point  $x$  d'un espace mesurable.

**Lemme 11.16.** Soient  $X$  un espace métrisable compact,  $\phi : X \rightarrow X$  une application continue et  $\epsilon > 0$ . Pour tout  $n \in \mathbb{N}$ , choisissons une partie  $E_n$  de  $X$  qui est  $\epsilon$ -séparée pour la distance  $d_n^\phi$ , et notons  $\nu_n = \frac{1}{\text{Card } E_n} \sum_{x \in E_n} \Delta_x$  et  $\mu_n = \frac{1}{n+1} \sum_{i=0}^n (\phi^i)_*(\nu_n)$ . Alors la suite  $(\mu_n)_{n \in \mathbb{N}}$  admet une sous-suite vaguement convergente vers une limite  $\mu$ . De plus, une telle mesure de probabilité  $\mu$  est invariante par  $\phi$  et nous avons

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln(\text{Card } E_n) \leq h_\mu(\phi) .$$

**Démonstration.** L'existence de  $\mu$  résulte de la compacité vague et métrisabilité vague de  $\text{Prob}(X)$  (voir la proposition 0.8). L'invariance par  $\phi$  de  $\mu$  résulte de la continuité vague de la mesure image (voir le lemme 0.7) et (sachant que  $\nu_n$  est une mesure de probabilité pour tout  $n \in \mathbb{N}$ ) d'un passage à la limite dans l'égalité

$$\phi_* \mu_n - \mu_n = \frac{1}{n+1} ((\phi^{n+1})_* \nu_n - \nu_n) .$$

D'après le lemme 11.15, il existe une partition mesurable  $\alpha$  de  $X$  telle que, pour tout  $A$  dans  $\alpha$ , nous avons  $\text{diam}(A) \leq \epsilon$  et  $\mu(\partial A) = 0$ . Pour tout  $n \in \mathbb{N}$ , posons  $\alpha_n = \bigvee_{i=0}^{n-1} \phi^{-i} \alpha$ . Comme chaque événement de la partition  $\alpha_n$  contient au plus un point de  $E_n$ , chacun d'eux étant de masse  $\frac{1}{\text{Card } E_n}$  pour la mesure  $\nu_n$ , nous avons l'égalité

$$H_{\nu_n}(\alpha_n) = - \sum_{A \in \alpha_n, A \cap E_n \neq \emptyset} \frac{1}{\text{Card } E_n} \ln \left( \frac{1}{\text{Card } E_n} \right) = \ln(\text{Card } E_n).$$

Soient  $0 \leq k < q \leq \frac{n}{2}$ . En notant  $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$  la partie entière (inférieure) de tout réel  $x$ , et en regroupant par paquets de  $q$  éléments consécutifs les entiers entre  $k$  et  $n-1$ , nous avons

$$\alpha_n = \left( \bigvee_{j=0}^{k-1} \phi^{-j} \alpha \right) \vee \left( \bigvee_{j=0}^{\lfloor \frac{n}{q} \rfloor - 2} \phi^{-jq-k} \alpha_q \right) \vee \left( \bigvee_{j=q\lfloor \frac{n}{q} \rfloor - q + k}^{n-1} \phi^{-j} \alpha \right).$$

Par la sous-additivité de l'entropie des partitions mesurables (voir la proposition 5.1 (5)), et puisqu'il y a au plus  $2q - k \leq 2q$  entiers entre  $q\lfloor \frac{n}{q} \rfloor - q + k$  et  $n-1$  nous avons alors

$$\ln(\text{Card } E_n) = H_{\nu_n}(\alpha_n) \leq \sum_{j=0}^{\lfloor \frac{n}{q} \rfloor - 2} H_{\nu_n}(\phi^{-k-qj} \alpha_q) + 3q \ln(\text{Card } \alpha).$$

Notons que  $H_{\nu_n}(\phi^{-i} \alpha_q) = H_{(\phi^i)_* \nu_n}(\alpha_q)$  pour tout  $i \in \mathbb{N}$ , par la naturalité de l'entropie des partitions (voir la proposition 5.1 (2)). En sommant la formule centrée précédente sur  $k \in \{0, \dots, q-1\}$ , puis en appliquant la concavité de l'entropie d'une partition en fonction de la mesure (voir la proposition 5.1 (6) pour les paramètres barycentriques constants égaux à  $\frac{1}{n+1}$ ), nous obtenons

$$q \log(\text{Card } E_n) \leq \sum_{i=0}^n H_{\nu_n}(\phi^{-i} \alpha_q) + 3q^2 \log(\text{Card } \alpha) \leq (n+1) H_{\mu_n}(\alpha_q) + 3q^2 \log(\text{Card } \alpha).$$

En divisant par  $nq$  et en faisant tendre  $n$  vers l'infini, nous obtenons

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log(\text{Card } E_n) \leq \frac{1}{q} \lim_{n \rightarrow \infty} H_{\mu_n}(\alpha_q) = \frac{1}{q} H_{\mu}(\alpha_q)$$

par le lemme 11.14. Nous faisons alors tendre  $q$  vers l'infini, ce qui donne

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log(\text{Card } E_n) \leq h_{\mu}(\phi, \alpha) \leq h_{\mu}(\phi).$$

Ceci démontre le lemme 11.16. □

**Fin de la démonstration du théorème 11.11.** Pour tout  $\epsilon > 0$ , appliquons le lemme précédent avec  $E_n$  de cardinal maximal parmi les parties finies de  $X$  qui sont  $\epsilon$ -séparées pour la distance  $d_n^{\phi}$ , de sorte que  $N'_{d_n^{\phi}}(\epsilon) = \text{Card } E_n$ . Pour tout  $\epsilon > 0$ , il existe donc un élément  $\mu_{\epsilon}$  dans  $\text{Prob}_{\phi}(X)$  tel que

$$\sup_{\mu \in \text{Prob}_{\phi}(X)} h_{\mu}(\phi) \geq h_{\mu_{\epsilon}}(\phi) \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \log N'_{d_n^{\phi}}(\epsilon).$$

D'après la proposition 11.5, le membre de droite tend vers  $h_{\text{top}}(\phi)$  quand  $\epsilon$  tend vers 0. Nous obtenons donc, comme souhaité, que  $\sup_{\mu \in \text{Prob}_{\phi}(X)} h_{\mu}(\phi) \geq h_{\text{top}}(\phi)$ . □

**Corollaire 11.17.** *Nous avons  $h_{\text{top}}(\phi) = h_{\text{top}}(\phi|_{\Omega(\phi)})$*

**Démonstration.** Par la proposition 7.3, le support de toute mesure finie invariante par  $\phi$  est contenu dans l'ensemble non errant  $\Omega(\phi)$  de  $\phi$ . Le résultat découle alors du principe variationnel.  $\square$

Soit  $(X, \phi)$  un système dynamique topologique, où  $X$  est un espace topologique localement compact. Une *mesure d'entropie maximale* de  $(X, \phi)$  est une mesure de probabilité borélienne  $\mu_{\text{max}}$  sur  $X$  invariante par  $\phi$  telle que

$$h_{\mu_{\text{max}}}(\phi) = \sup_{\mu \in \text{Prob}_{\phi}(X)} h_{\mu}(\phi).$$

Le problème de l'existence d'une mesure d'entropie maximale (et plus généralement des mesures d'équilibre), le problème de l'unicité (ou de la compréhension de l'espace des mesures d'entropie maximale) lorsqu'il y a existence, et les problèmes de constructions effectives lorsqu'il y a existence et unicité, sont des problèmes fondamentaux en théorie ergodique, et encore largement ouverts en absence d'hypothèses de compacité ou d'hyperbolicité uniforme (voir par exemple [Bow2, Man, New, BriD, BuzS, Buz1, Bur, Buz2, BuzCS], mais ce n'est qu'un échantillon aléatoire).

Le résultat suivant, immédiat par la proposition 5.3 (5) et la proposition 11.4, dit que ces problèmes d'existence et d'unicité sont invariants par conjugaison.

**Proposition 11.18.** *Soient  $(X, \phi)$  et  $(X', \phi')$  deux systèmes dynamiques topologiques conjugués par un homéomorphisme  $h : X \rightarrow X'$ , avec  $X$  et  $X'$  compacts. L'application  $h_*$  de mesure image par  $h$  est une bijection de l'ensemble des mesures d'entropie maximale de  $(X, \phi)$  dans l'ensemble des mesures d'entropie maximale de  $(X', \phi')$ .  $\square$*

En particulier, le système dynamique  $(X, \phi)$  admet au moins une mesure d'entropie maximale (respectivement admet une unique mesure d'entropie maximale) si et seulement s'il en est de même pour  $(X', \phi')$ .

**Corollaire 11.19.** *Un système dynamique topologique expansif  $(X, \phi)$ , où  $X$  est un espace métrique compact, admet au moins une mesure d'entropie maximale  $\mu_{\text{max}}$  telle que  $h_{\mu_{\text{max}}}(\phi) = h_{\text{top}}(\phi)$ .*

**Démonstration.** Par la deuxième affirmation du corollaire 11.10, il existe  $\epsilon_0 > 0$  tel que

$$h_{\text{top}}(\phi) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log N'_{d_n^{\phi}}(\epsilon_0).$$

En appliquant le lemme 11.16 avec  $E_n$  de cardinal maximal parmi les parties finies de  $X$  qui sont  $\epsilon_0$ -séparées pour la distance  $d_n^{\phi}$ , il existe donc un élément  $\mu$  dans  $\text{Prob}_{\phi}(X)$  tel que  $h_{\mu}(\phi) \geq h_{\text{top}}(\phi)$ . L'inégalité inverse étant vérifiée par le sens facile du principe variationnel, le résultat en découle.  $\square$

**Exemples.** (1) Pour tout  $N \geq 2$ , la mesure de Lebesgue est l'unique mesure d'entropie maximale pour les applications  $t \mapsto Nt$  de  $\mathbb{R}/\mathbb{Z}$  dans lui-même (voir l'exercice E.71).

(2) Si  $(X, \phi)$  est un décalage de Bernoulli sur un alphabet fini  $\mathcal{A}$ , si  $\nu$  est la mesure d'équiprobabilité sur  $\mathcal{A}$ , alors la mesure produit  $\nu^{\mathbb{N}}$  dans le cas unilatère et  $\nu^{\mathbb{Z}}$  dans le cas

bilatère est l'unique mesure d'entropie maximale (voir l'exercice E.72). Nous généraliserons ceci dans la partie 11.3 pour les sous-décalages de type fini.

(3) Nous montrerons dans la partie 11.4 que la mesure de Haar sur le tore  $\mathbb{T}^N$  est l'unique mesure d'entropie maximale pour un automorphisme linéaire hyperbolique du tore.

### 11.3 Entropie topologique des systèmes dynamiques symboliques

Les références recommandées pour cette partie sont [PY, Cou].

Une description combinatoire ou symbolique d'un système dynamique est parfois souhaitable (même si elle n'existe pas toujours). Un *codage* d'un système dynamique topologique  $(X, \phi)$  consiste à repérer l'orbite d'un point  $x$  de  $X$  à partir d'une décomposition finie  $X = \bigcup_{i \in \mathcal{A}} R_i$  en fermés d'intérieurs disjoints. Nous dirons qu'une suite  $\omega = (\omega_n)_n$  de  $\mathcal{A}^{\mathbb{N}}$  (ou de  $\mathcal{A}^{\mathbb{Z}}$  si  $\phi$  est inversible) *code* le point  $x$  si, à chaque temps  $n$ , le point  $\phi^n(x)$  appartient à  $R_{\omega_n}$ , autrement dit si  $x$  appartient à  $\bigcap_n \phi^{-n}(R_{\omega_n})$ .

Choisir un bon codage, c'est choisir une décomposition de sorte que les intersections  $\bigcap_n \phi^{-n}(R_{\omega_n})$  contiennent au plus un point et telles que l'on puisse facilement décrire les suites  $\omega$  pour lesquelles cette intersection est non vide. Le système dynamique  $(X, \phi)$  apparaît alors comme une semi-conjugaison d'un système de Bernoulli ou d'un sous-espace invariant d'un système de Bernoulli. Le premier but de cette partie est d'introduire une classe de tels sous-systèmes qui seront bien adaptés aux divers calculs, appelés les sous-décalages de type fini.

Le prototype d'un codage est l'écriture décimale d'un réel  $x$  de  $[0, 1[$ . Ceci code le système dynamique  $(X = \mathbb{R}/\mathbb{Z}, \phi : x \mapsto 10x)$ , en prenant pour alphabet fini  $\mathcal{A} = \{0, 1, \dots, 9\}$  et pour parties les  $R_i = [\frac{i}{10}, \frac{i+1}{10}]$  pour  $i$  dans  $\mathcal{A}$ . L'application de codage (non injective) est l'application de  $\mathcal{A}^{\mathbb{N}}$  dans  $\mathbb{R}/\mathbb{Z}$  définie par

$$(a_n)_{n \in \mathbb{N}} \mapsto \sum_{n \in \mathbb{N}} \frac{a_n}{10^{n+1}}.$$

Un autre exemple fondamental de codage est celui que nous avons effectué pour l'étude du fer à cheval sur l'attracteur de Smale (voir la partie 10.5). Dans cet exemple, la semi-conjugaison était une conjugaison, et il n'y avait pas besoin de passer à un sous-système du système de Bernoulli, mais cette situation est loin d'être générale.

La discussion ci-dessus justifie l'importance des sous-décalages de type fini. Cette partie leur est entièrement consacré. Nous calculerons leur entropie topologique. Nous montrerons que, pour les sous-décalages de type fini vérifiant une propriété simple d'irréductibilité, il existe une unique mesure d'entropie maximale, et que cette mesure reflète la répartition statistique des orbites périodiques.

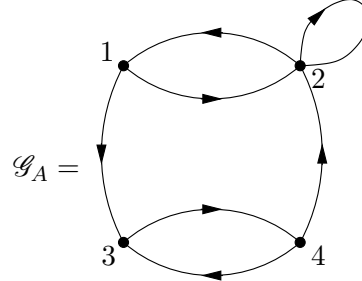
Nous ne considérerons que les codages avec alphabet fini dans ces notes. Mais il est parfois utile, en particulier en l'absence de compacité de l'espace des phases (voir par exemple [BrPP, Chap.5]), de considérer des codages par des alphabets dénombrables. Nous renvoyons pour cela par exemple aux références [Gur, Kit, Sar1, Sar4, Sar5, Woe].

#### 11.3.1 Sous-décalages de type fini

Soient  $\mathcal{A}$  un ensemble fini de cardinal au moins 2, appelé un *alphabet*, muni de la topologie discrète, et  $A = (A_{\alpha\beta})_{\alpha, \beta \in \mathcal{A}}$  une matrice carrée de coefficients égaux à 0 ou 1. Se

donner une telle matrice revient à se donner un graphe orienté  $\mathcal{G}_A$  d'ensemble de sommets  $\mathcal{A}$ , tel qu'il existe une arête allant du sommet  $\alpha \in \mathcal{A}$  au sommet  $\beta \in \mathcal{A}$  si et seulement si  $A_{\alpha,\beta} = 1$ . La matrice  $A$  est alors appelée la *matrice d'incidence*<sup>242</sup> du graphe orienté  $\mathcal{G}_A$ .

**Exemple.** 
$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$



Nous appellerons *sous-décalage de type fini*<sup>243</sup> bilatère de matrice d'incidence  $A$  (ou de graphe  $\mathcal{G}_A$ ) le système dynamique topologique inversible (à temps discret)  $(\Sigma_A, \sigma_A)$  où

- $\Sigma_A$  est la partie fermée donc compacte de  $\Sigma = \mathcal{A}^{\mathbb{Z}}$  définie par

$$\Sigma_A = \{ \omega \in \mathcal{A}^{\mathbb{Z}} : \forall n \in \mathbb{Z}, A_{\omega_n, \omega_{n+1}} = 1 \},$$

- $\sigma_A$  est la restriction du décalage de Bernoulli bilatère  $\sigma$  à  $\Sigma_A$ .

Nous munissons  $\Sigma_A$  de la distance induite  $d$  et de la topologie induite par celles de  $\Sigma$  (voir la formule (4)).

Nous définissons de même le *sous-décalage de type fini*<sup>244</sup> unilatère de matrice d'incidence  $A$  (ou de graphe  $\mathcal{G}_A$ ) comme le système dynamique topologique (à temps discret)  $(\Sigma_{A,+}, \sigma_{A,+})$  obtenu en remplaçant  $\mathbb{Z}$  par  $\mathbb{N}$  (et donc  $(\Sigma, \sigma)$  par  $(\Sigma_+, \sigma_+)$ ) ci-dessus.

Si  $A$  est la matrice dont tous les coefficients sont égaux à 1, ou, de manière équivalente, si  $\mathcal{G}$  est le graphe orienté complet, alors nous retrouvons les systèmes de Bernoulli :  $(\Sigma_A, \sigma_A) = (\Sigma, \sigma)$  et  $(\Sigma_{A,+}, \sigma_{A,+}) = (\Sigma_+, \sigma_+)$ .

Si  $A$  est une matrice d'incidence, nous avons déjà vu dans la partie 6.1 que sauf cas particulier, il existe de très nombreuses mesures invariantes sur les sous-décalages de type fini  $(\Sigma_A, \sigma_A)$  et  $(\Sigma_{A,+}, \sigma_{A,+})$ .

**Exercice E.66.** Soit  $A = (A_{\alpha,\beta})_{\alpha,\beta \in \mathcal{A}}$  une matrice d'incidence. Montrer que les assertions suivantes sont équivalentes :

- (1) la matrice  $A$  est irréductible
- (2) le graphe orienté  $\mathcal{G}_A$  est connexe (en tant que graphe orienté),
- (3)  $A$  n'a pas de ligne ou de colonne nulle et le système dynamique topologique inversible  $(\Sigma_A, \sigma_A)$  est positivement transitif,
- (4)  $A$  n'a pas de ligne ou de colonne nulle et le système dynamique topologique inversible  $(\Sigma_A, \sigma_A)$  admet une orbite dense,
- (5)  $A$  n'a pas de ligne ou de colonne nulle et le système dynamique topologique  $(\Sigma_{A,+}, \sigma_{A,+})$  est positivement transitif,
- (6)  $A$  n'a pas de ligne ou de colonne nulle et le système dynamique topologique  $(\Sigma_{A,+}, \sigma_{A,+})$  admet un point dont l'ensemble  $\omega$ -limite est égal à  $\Sigma_{A,+}$ .

242. Certains ouvrages, dont [Kit], disent *matrice de transition*, mais nous préférons réserver cette terminologie aux matrices de transition des chaînes de Markov, voir la partie 6.1.

243. Certaines références disent *système de Markov*, mais nous préférons n'utiliser la terminologie markovienne que pour les aspects de théorie de la mesure et de probabilités des chaînes de Markov.

244. Voir la note de bas de page précédente.

Si  $A$  est irréductible, montrer que  $\Sigma_A$  est ou bien fini et réduit à une orbite périodique, ou bien un espace de Cantor.

**Exercice E.67.** Montrer qu'un sous-décalage de type fini est expansif, et d'entropie topologique finie.

Nous allons calculer l'entropie topologique d'un sous-décalage de type fini.

L'entropie topologique d'un sous-décalage de type fini est donnée par le résultat suivant.

**Proposition 11.20.** Soit  $\rho(A) = \lim_{n \rightarrow \infty} \|A^n\|^{\frac{1}{n}}$  le rayon spectral de la matrice  $A$  (qui ne dépend pas du choix d'une norme matricielle<sup>245</sup>  $\|\cdot\|$  sur l'algèbre  $\mathcal{M}_{\mathcal{A}}(\mathbb{R})$ ), alors

$$h_{\text{top}}(\sigma_A) = h_{\text{top}}(\sigma_{A,+}) = \ln \rho(A).$$

Si  $A$  est apériodique, alors  $h_{\text{top}}(\sigma_A) = h_{\text{top}}(\sigma_{A,+}) = \ln \lambda_{\max}(A)$ .

**Démonstration.** La seconde assertion découle de la première et du théorème de Perron-Frobenius 6.1. Nous montrons le cas bilatère (c'est-à-dire l'égalité  $h_{\text{top}}(\sigma_A) = \ln \rho(A)$ ), le cas unilatère se traite de manière similaire. Nous aurons aussi besoin des deux assertions du lemme de comptage ci-dessous ultérieurement.

**Lemme 11.21.**

(1) Pour tout  $\alpha$  dans  $\mathcal{A}$ , notons  $[\alpha]_0$  le 1-cylindre  $\{\omega \in \Sigma_A : \omega_0 = \alpha\}$ . Alors pour tous les  $n \in \mathbb{N}$  et  $a, b \in \mathcal{A}$ , le cardinal de l'ensemble

$$\{(\alpha_0, \dots, \alpha_n) \in \mathcal{A}^{n+1} : [\alpha_0]_0 \cap \sigma_A^{-1}([\alpha_1]_0) \cap \dots \cap \sigma_A^{-n}([\alpha_n]_0) \neq \emptyset, \alpha_0 = a, \alpha_n = b\}$$

est égal au coefficient  $(A^n)_{a,b}$  de la matrice  $A^n$ .

(2) Pour tout  $N \in \mathbb{N}$ , si  $P_N = \{\omega \in \Sigma_A : \sigma_A^N \omega = \omega\}$ , alors

$$\text{Card } P_N = \text{tr } A^N.$$

**Démonstration.** (1) Cet ensemble est aussi l'ensemble

$$\{(\alpha_0, \dots, \alpha_n) \in \mathcal{A}^{n+1} : \alpha_0 = a, \alpha_n = b, \text{ et } \forall k = 0, \dots, n-1, A_{\alpha_k, \alpha_{k+1}} = 1\}.$$

Son cardinal est donc le nombre de chemins joignant  $a$  à  $b$  de longueur  $n$  dans le graphe orienté  $\mathcal{G}_A$ . Ce nombre est exactement le coefficient  $(A^n)_{a,b}$  de la matrice  $A^n$ .

(2) Ceci découle du fait que  $\text{tr } A^N = \sum_{a \in \mathcal{A}} (A^N)_{a,a}$ . □

Pour tout  $n \in \mathbb{N}$ , appelons dans cette démonstration  $(2n+1)$ -cylindre toute partie de  $\Sigma_A$  de la forme  $\{\omega \in \Sigma_A : \omega_{-n} = a_{-n}, \dots, \omega_n = a_n\}$  pour  $a_{-n}, \dots, a_n \in \mathcal{A}$ . Deux éléments qui appartiennent à un même  $(2n+1)$ -cylindre ont les mêmes coefficients d'indice entre  $-n$  et  $n$ . Donc par la définition (4) de la distance  $d$  sur  $\Sigma_A$ , le diamètre d'un  $(2n+1)$ -cylindre est au plus  $e^{-n}$ .

Par la définition de la topologie produit, les  $(2n+1)$ -cylindres sont des ouverts de  $\Sigma_A$ . Le recouvrement ouvert  $\mathcal{U}$  de  $\Sigma_A$  par les 1-cylindres est générateur pour  $\sigma_A$ , car pour tout

<sup>245.</sup> c'est-à-dire d'une norme  $\|\cdot\|$  sur l'algèbre  $\mathcal{M}_{\mathcal{A}}(\mathbb{R})$  des matrices carrées de lignes et colonnes indexées par  $\mathcal{A}$  vérifiant, pour toutes les matrices  $A, B \in \mathcal{M}_{\mathcal{A}}(\mathbb{R})$ , l'inégalité  $\|AB\| \leq \|A\| \|B\|$

$n \in \mathbb{N}$ , le recouvrement ouvert  $\bigvee_{i=-n}^n \sigma_A^{-i} \mathcal{U}$  est le recouvrement par les  $(2n+1)$ -cylindres, donc son diamètre est au plus  $e^{-n}$ . Deux  $(2n+1)$ -cylindres sont égaux ou disjoints, donc  $N(\bigvee_{i=-n}^n \sigma_A^{-i} \mathcal{U})$  est égal au nombre de  $(2n+1)$ -cylindres non vides.

Par l'équivalence des normes en dimension finie, il existe  $c \geq 1$  telle que, pour toute matrice  $B \in \mathbb{R}^{\mathcal{A}^2}$ , nous avons

$$\frac{1}{c} \|B\| \leq \sum_{\alpha, \beta \in \mathcal{A}} |B_{\alpha, \beta}| \leq c \|B\|. \quad (134)$$

Par le théorème 11.9, par la formule (133), par le lemme 11.21 (1) et par la formule (134), nous avons donc

$$\begin{aligned} h_{\text{top}}(\sigma_A) &= h(\sigma_A, \mathcal{U}) = \lim_{n \rightarrow \infty} \frac{1}{2n+1} \ln N\left(\bigvee_{i=-n}^n \sigma_A^{-i} \mathcal{U}\right) = \lim_{n \rightarrow \infty} \frac{1}{2n+1} \ln \sum_{a, b \in \mathcal{A}} (A^{2n+1})_{a, b} \\ &= \lim_{n \rightarrow \infty} \frac{1}{2n+1} \ln \|A^{2n+1}\| = \ln \rho(A). \quad \square \end{aligned}$$

### 11.3.2 Problématique de l'équidistribution

Soit  $X$  un espace topologique localement compact, de tribu borélienne  $\mathcal{B}$ , muni d'une mesure (borélienne positive) régulière  $\mu$ . Notons  $(\lambda_N)_{N \in \mathbb{N}}$  une suite de réels strictement positifs et  $\mathcal{F} = (F_N)_{N \in \mathbb{N}}$  une suite d'ensembles dénombrables  $F_N$  de parties fermées  $f$  de  $X$  munies d'une mesure régulière  $\mu_f$  de support  $f$ , tels que la mesure

$$\mu_N = \frac{1}{\lambda_N} \sum_{f \in F_N} \mu_f$$

soit régulière. Ceci est en particulier le cas si  $F_N$  est fini.

Par exemple,  $X$  peut être compact,  $\mu$  une mesure de probabilité, et  $F_N$  un ensemble fini de singletons  $f = \{x\}$  de  $X$ , que nous identifions à un ensemble fini de points  $x$  de  $X$ , munis de leur masse de Dirac unité  $\mu_f = \Delta_x$ , avec  $\lambda_N = \text{Card } F_N$ .

Soit  $\mathcal{B}_\mu$  l'ensemble des boréliens  $B \in \mathcal{B}$  d'adhérence compacte tels que  $\mu(\partial B) = 0$ .<sup>246</sup> Des résultats classiques de théorie de la mesure (voir par exemple [Coh] ou le lemme de Portmanteau de [Bil]) montrent le résultat suivant.

**Proposition 11.22.** *Les conditions suivantes sont équivalentes :*

- (1) *la suite de mesures  $(\mu_N)_{N \in \mathbb{N}}$  converge vaguement vers la mesure  $\mu$  (c'est-à-dire (voir la partie 0.2) que pour toute application continue à support compact  $f \in C_c^0(X)$ , nous avons  $\lim_{N \rightarrow +\infty} \mu_N(f) = \mu(f)$ ),*
- (2) *pour tout  $B \in \mathcal{B}_\mu$ , nous avons  $\lim_{N \rightarrow +\infty} \mu_N(B) = \mu(B)$ ,*
- (3) *il existe une partie  $\mathcal{B}'$  de  $\mathcal{B}_\mu$  qui engendre la  $\sigma$ -algèbre  $\mathcal{B}$  telle que pour tout  $B \in \mathcal{B}'$ , nous ayons  $\lim_{N \rightarrow +\infty} \mu_N(B) = \mu(B)$ . □*

Si l'une de ces assertions est vérifiée, nous dirons alors que la suite  $\mathcal{F}$  s'équidistribue dans  $X$  vers la mesure  $\mu$  à l'échelle  $(\lambda_N)_{N \in \mathbb{N}}$ .

246. Il est important de ne pas oublier cette condition.

Lorsque  $X$  est compact, si  $F_N$  est fini et  $\lambda_N = \text{Card } F_N$ , si les mesures  $\mu_f$  pour  $f \in F_N$  sont des mesures de probabilité, la mesure  $\mu$ , lorsqu'elle existe, doit alors être une mesure de probabilité. Mais lorsque  $X$  n'est pas compact, il peut apparaître des phénomènes intéressants de perte de masse à l'infini de la suite  $(\nu_N)_{N \in \mathbb{N}}$  : la mesure  $\mu$  limite pour la topologie vague, lorsqu'elle existe, n'est pas forcément une mesure de probabilité si  $X$  n'est pas compact.

Donnons des exemples de parties  $\mathcal{B}'$  comme dans l'assertion (3) ci-dessus. Lorsque  $(X, \mu)$  est la droite réelle munie de la mesure de Lebesgue, nous pouvons prendre pour  $\mathcal{B}'$  l'ensemble des intervalles bornés dans  $\mathbb{R}$ , voire l'ensemble des intervalles bornés dont les deux extrémités appartiennent à une partie dense fixée de  $\mathbb{R}$ . Lorsque  $(X, \mu)$  est l'espace euclidien  $\mathbb{R}^n$  muni de la mesure de Lebesgue, nous pouvons prendre pour  $\mathcal{B}'$  l'ensemble des cubes dans  $\mathbb{R}^n$ . Lorsque  $X$  est l'espace de phases d'un sous-décalage de type fini, nous pouvons prendre pour  $\mathcal{B}'$  l'ensemble des cylindres, et ceci pour toutes les mesures boréliennes, la frontière d'un cylindre étant vide.

Lorsque  $F_N$  est un ensemble fini de points (munis de leur masse de Dirac unité) de  $X$  et  $\lambda_N = \text{Card } F_N$ , dire que  $\mathcal{F}$  s'équidistribue dans  $X$  vers la mesure  $\mu$  (c'est-à-dire que  $(\mu_N)_{N \in \mathbb{N}}$  converge vaguement vers la mesure  $\mu$ ) est alors équivalent à dire que pour tout élément<sup>247</sup>  $B \in \mathcal{B}_\mu$  (ou, de manière équivalente, pour tout élément  $B$  d'une partie  $\mathcal{B}'$  donnée de  $\mathcal{B}_\mu$  qui engendre la  $\sigma$ -algèbre  $\mathcal{B}$ ), la proportion de points de  $F_N$  dans  $B$  est asymptotiquement égale à la mesure de  $B$  pour  $\mu$  :

$$\lim_{N \rightarrow +\infty} \frac{\text{Card}(B \cap F_N)}{\text{Card } F_N} = \mu(B).$$

**Exemples.** (1) Par les propriétés de l'intégrale de Riemann, la suite

$$\mathcal{F} = \left( F_N = \left\{ \frac{k}{N+1} : 0 \leq k \leq N+1 \right\} \right)_{N \in \mathbb{N}}$$

de parties finies de  $[0, 1]$  s'équidistribue dans l'intervalle  $[0, 1]$  vers la mesure de Lebesgue  $\text{Leb}_{[0,1]}$ .

(2) Soit  $\alpha \in \mathbb{R} - \mathbb{Q}$  un nombre réel irrationnel. Pour tout réel  $t \in \mathbb{R}$ , notons de manière usuelle  $\{t\} = t - \lfloor t \rfloor \in [0, 1[$  la *partie fractionnaire* de  $t$ . Par l'unique ergodicité de la rotation  $R_\alpha$  sur  $\mathbb{R}/\mathbb{Z}$  (et la version adaptée du théorème ergodique de Birkhoff, voir la proposition 4.2), la suite

$$\mathcal{F} = \left( F_N = \{ \{n\alpha\} : 0 \leq n \leq N \} \right)_{N \in \mathbb{N}}$$

de parties finies de  $[0, 1]$  s'équidistribue dans l'intervalle  $[0, 1]$  vers la mesure de Lebesgue  $\text{Leb}_{[0,1]}$ . Plus généralement, si  $P$  est un polynôme réel en une variable dont au moins l'un des coefficients est irrationnel, alors la suite des  $F_N = \{ \{P(n)\} : 0 \leq n \leq N \}$  s'équidistribue dans  $[0, 1]$  vers  $\text{Leb}_{[0,1]}$ . Ce résultat, originellement dû à Weyl, se démontre aussi par les méthodes de séries de Fourier (voir la partie 4).<sup>248</sup>

247. attention à ne pas oublier de vérifier l'hypothèse que  $\mu(\partial B) = 0$

248. Le *critère de Weyl* (voir par exemple [IK, §21], ainsi que pour de nombreux autres exemples d'équidistribution de suites arithmétiques) dit qu'une suite réelle  $(x_n)_{n \in \mathbb{N}}$  s'équidistribue modulo 1 (c'est-à-dire la suite des  $F_N = \{ \{x_n\} : 0 \leq n \leq N \}$  s'équidistribue dans  $[0, 1]$  vers  $\text{Leb}_{[0,1]}$ ) si et seulement si  $\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=0}^N e^{2\pi i \ell x_n} = 0$  pour tout  $\ell \in \mathbb{Z} - \{0\}$ . Ceci découle de la densité du sous-espace vectoriel complexe engendré par les fonctions trigonométriques dans l'espace de Banach des fonctions continues de  $[0, 1]$  dans  $\mathbb{C}$  pour la norme uniforme, voir la proposition 4.1 (1).



(3) Soit  $\alpha \in ]0, 1[$ . Puisque  $(n+1)^\alpha - n^\alpha \sim \alpha/n^{1-\alpha}$  quand  $n \rightarrow +\infty$ , la suite

$$\mathcal{F} = \left( F_N = \{ \{n^\alpha\} : 0 \leq n \leq N \} \right)_{N \in \mathbb{N}}$$

de parties finies de  $[0, 1]$  s'équidistribue dans l'intervalle  $[0, 1]$  vers la mesure de Lebesgue  $\text{Leb}_{[0,1]}$ .

(4) La suite de Farey

$$\mathcal{F} = \left( F_N = \left\{ \frac{p}{q} : 0 \leq p \leq q, 0 < q \leq N, (p, q) = 1 \right\} \right)_{N \in \mathbb{N}}$$

de parties finies de  $[0, 1]$  s'équidistribue dans l'intervalle  $[0, 1]$  vers la mesure de Lebesgue  $\text{Leb}_{[0,1]}$  (voir [Nevi], ainsi que [PaP1] pour une démonstration géométrique et [BrPP, Chap. 16] pour une version en caractéristique positive).

(5) Le dessin en couverture de ces notes représente l'équidistribution de points dans  $\mathbb{Q}(e^{\frac{2i\pi}{3}})$  dans  $\mathbb{C}$ , voir par exemple [PaP2] pour la description de la famille  $\mathcal{F}$  considérée.

Nous allons nous intéresser à des problèmes d'équidistribution de points périodiques de systèmes dynamiques, ou d'autres familles  $\mathcal{F}$  construites par des moyens dynamiques ou arithmétiques, dans le reste de ces notes.

### 11.3.3 Mesures d'entropie maximale des sous-décalages de type fini

Le but de cette partie est de montrer qu'il existe une et une seule mesure d'entropie maximale pour tout sous-décalage de type fini de matrice d'incidence apériodique, de donner une construction précise de cette mesure, et d'en déduire l'équidistribution des points périodiques dans un sous-décalage de type fini.

Notons encore  $\mathcal{A}$  un alphabet fini de cardinal au moins 2, et  $A = (A_{\alpha\beta})_{\alpha, \beta \in \mathcal{A}}$  une matrice d'incidence (de coefficients égaux à 0 ou 1). Supposons que  $A$  soit apériodique. Avec les notations du théorème 6.1 de Perron-Frobenius, notons  $\lambda = \lambda_{\max}(A)$ ,  $v = (v_\alpha)_{\alpha \in \mathcal{A}}$  un vecteur propre de  ${}^tA$  pour la valeur propre  $\lambda$  à coefficients strictement positifs, et  $f = (f_\alpha)_{\alpha \in \mathcal{A}}$  un vecteur propre de  $A$  pour la valeur propre  $\lambda$  à coefficients strictement positifs, normalisés<sup>249</sup> de sorte que

$$\sum_{\alpha \in \mathcal{A}} v_\alpha f_\alpha = 1. \quad (135)$$

Nous avons donc, pour tout  $\alpha$  dans  $\mathcal{A}$ ,

$$v_\alpha > 0, \quad f_\alpha > 0, \quad \sum_{\beta \in \mathcal{A}} v_\beta A_{\beta\alpha} = \lambda v_\alpha, \quad \text{et} \quad \sum_{\beta \in \mathcal{A}} A_{\alpha\beta} f_\beta = \lambda f_\alpha. \quad (136)$$

Notons alors  $\pi^\bullet = (\pi_{\alpha\beta}^\bullet)_{\alpha, \beta \in \mathcal{A}}$  la matrice de coefficients

$$\pi_{\alpha\beta}^\bullet = \frac{A_{\alpha\beta} f_\beta}{\lambda f_\alpha}. \quad (137)$$

---

249. Il reste encore plusieurs choix possibles, puisqu'il est possible de multiplier  $v$  et de diviser  $f$  par un même  $\lambda > 0$ , mais ce qui suit ne dépend pas d'un tel choix.

Il est immédiat de vérifier par cette formule (137) que  $\pi^\bullet$  est une matrice de transition apériodique, dont la matrice d'incidence associée est égale à  $A$ . La mesure de Markov associée

$$\mu = \mathbb{P}_{\pi^\bullet} \quad (138)$$

sur  $\Sigma_A$ , respectivement

$$\mu_+ = \mathbb{P}_{\pi^\bullet, +} \quad (139)$$

sur  $\Sigma_{A,+}$ , s'appelle la *mesure de Parry* du sous-décalage de type fini bilatère  $(\Sigma_A, \sigma_A)$ , respectivement unilatère  $(\Sigma_{A,+}, \sigma_{A,+})$ .

**Exemple.** Lorsque  $A$  est une matrice carrée de coefficients tous égaux à 1, alors  $\lambda$  vaut  $q = \text{Card } \mathcal{A}$ ,  $v_\alpha = f_\alpha = \frac{1}{\sqrt{q}}$  pour tout  $\alpha$  dans  $\mathcal{A}$ , et  $\pi^\bullet$  est la matrice carrée de coefficients tous égaux à  $\frac{1}{q}$ . Le sous-décalage de type fini  $(\Sigma_A, \sigma_A)$  est alors égal au système de Bernoulli  $(\Sigma, \sigma)$  et la mesure de Parry de  $(\Sigma_A, \sigma_A) = (\Sigma, \sigma)$  est alors la mesure produit  $\nu_0^{\mathbb{Z}}$ , pour  $\nu_0$  la mesure d'équiprobabilité sur  $\mathcal{A}$ .

**Remarque 11.23.** L'unique (par le lemme 6.2 (2)) mesure  $\pi^\bullet$ -stationnaire est

$$\nu^{\pi^\bullet} = (v_\alpha f_\alpha)_{\alpha \in \mathcal{A}},$$

car  $\nu^{\pi^\bullet}$  est un vecteur de coordonnées strictement positives dont la somme est égale à 1 par la formule (135), et  $\nu^{\pi^\bullet} \pi^\bullet = \nu^{\pi^\bullet}$  par les formules (136) et (137).

Le résultat suivant montre l'existence et l'unicité d'une mesure d'entropie maximale sur tout sous-décalage de type fini de matrice d'incidence apériodique, égale à la mesure de Parry que nous venons de construire explicitement, et montre l'équidistribution des points périodiques vers la mesure d'entropie maximale.

**Théorème 11.24.** Soient  $(\Sigma_A, \sigma_A)$  un sous-décalage de type fini bilatère, de matrice d'incidence  $A$  apériodique, et  $\mu$  la mesure de Parry sur  $\Sigma_A$ . Pour tout  $N \in \mathbb{N}$ , notons

$$P_N = \{\omega \in \Sigma_A : \sigma_A^N \omega = \omega\}$$

l'ensemble des points périodiques de  $\sigma_A$  de période divisant  $N$ , et  $\mu_N = \frac{1}{\text{Card } P_N} \sum_{\omega \in P_N} \Delta_\omega$ , où  $\Delta_x$  est la masse de Dirac unité en  $x$ .

- a) Nous avons  $h_{\text{top}}(\sigma_A) = h_\mu(\sigma_A)$ .
- b) La seule mesure de probabilité  $\sigma_A$ -invariante  $\mu'$  sur  $\Sigma_A$  telle que  $h_\nu(\sigma_A) = h_{\text{top}}(\sigma_A)$  est  $\mu' = \mu$ .
- c) La famille  $\mathcal{F} = (P_N)_{N \in \mathbb{N}}$  s'équidistribue dans  $\Sigma_A$  vers la mesure d'entropie maximale : la suite des mesures de probabilité  $(\mu_N)_{N \in \mathbb{N}}$  converge vaguement vers  $\mu$ .

Une démonstration analogue donne que si  $(\Sigma_{A,+}, \sigma_{A,+})$  est un sous-décalage de type fini unilatère, de matrice d'incidence  $A$  apériodique, et si  $\mu_+$  est sa mesure de Parry, alors

- $\mu_+$  est l'unique mesure d'entropie maximale de  $(\Sigma_{A,+}, \sigma_{A,+})$ ,
- $h_{\text{top}}(\sigma_{A,+}) = h_{\mu_+}(\sigma_{A,+})$ , et
- la famille  $\mathcal{F} = (P_N)_{N \in \mathbb{N}}$  des ensembles  $P_N^+ = \{\omega \in \Sigma_{A,+} : \sigma_{A,+}^N \omega = \omega\}$  des points périodiques de période divisant  $N$  de  $\sigma_{A,+}$  s'équidistribue vers l'unique mesure d'entropie maximale de  $(\Sigma_{A,+}, \sigma_{A,+})$ .

En particulier, l'unique mesure d'entropie maximale du décalage de Bernoulli  $\sigma$  sur  $\mathcal{A}^{\mathbb{Z}}$ , pour  $\mathcal{A}$  un alphabet fini de cardinal au moins 2, est la mesure produit  $\nu_0^{\mathbb{Z}}$ , pour  $\nu_0$  l'équiprobabilité sur  $\mathcal{A}$  (voir aussi l'exercice E.72).

**Démonstration.** a) La proposition 6.4 appliquée à  $\pi = \pi^\bullet$  et à  $\nu = \nu^{\pi^\bullet}$ , la remarque 11.23 et la formule (137) donnent

$$h_\mu(\sigma_A) = \sum_{\alpha, \beta \in \mathcal{A}} \frac{v_\alpha A_{\alpha\beta} f_\beta}{\lambda} (-\ln(A_{\alpha\beta} f_\beta) + \ln f_\alpha + \ln \lambda).$$

Remarquons que  $A_{\alpha\beta} \ln A_{\alpha\beta} = 0$  car  $A_{\alpha\beta} \in \{0, 1\}$ . Donc par les formules (136) et (135), nous avons

$$\begin{aligned} \sum_{\alpha, \beta} -\ln(A_{\alpha\beta} f_\beta) \frac{v_\alpha A_{\alpha\beta} f_\beta}{\lambda} &= -\frac{1}{\lambda} \sum_{\alpha, \beta} \ln(f_\beta) v_\alpha A_{\alpha\beta} f_\beta = -\sum_{\beta} \ln(f_\beta) v_\beta f_\beta, \\ \sum_{\alpha, \beta} \ln(f_\alpha) \frac{v_\alpha A_{\alpha\beta} f_\beta}{\lambda} &= \sum_{\alpha} \ln(f_\alpha) v_\alpha f_\alpha, \\ \sum_{\alpha, \beta} \ln \lambda \frac{v_\alpha A_{\alpha\beta} f_\beta}{\lambda} &= \ln \lambda \sum_{\alpha} v_\alpha f_\alpha = \ln \lambda. \end{aligned}$$

Comme  $h_{\text{top}}(\sigma_A) = \ln \lambda$  par la dernière affirmation de la proposition 11.20, le résultat en découle. Donc la mesure de Parry est une mesure d'entropie maximale du système dynamique topologique  $(\Sigma_A, \sigma_A)$ .

b) Soit  $\mu'$  une autre mesure de probabilité borélienne  $\sigma_A$ -invariante d'entropie maximale sur  $\Sigma_A$ , donc telle que  $h_{\mu'}(\sigma_A) = h_{\text{top}}(\sigma_A) = \ln \lambda$ .

Si  $\mu'$  est absolument continue par rapport à la mesure de Parry  $\mu$ , alors il existe, d'après le théorème de Radon-Nikodym, une fonction  $\psi$  positive ou nulle dans  $\mathbb{L}^1(\Sigma_A, \mu)$  telle que  $\mu' = \psi\mu$ . Cette fonction est  $\sigma_A$ -invariante  $\mu$ -presque partout. Comme  $\mu$  est ergodique (car mélangeante, voir la proposition 6.5), et puisque  $\mu$  et  $\mu'$  sont des mesures de probabilité, nous avons  $\psi = 1$  (au sens de l'égalité dans  $\mathbb{L}^1(\Sigma_A, \mu)$ ) et  $\mu' = \mu$ , ce qui est exactement ce que nous cherchons à démontrer.

Supposons donc par l'absurde que  $\mu'$  ne soit pas absolument continue par rapport à  $\mu$ . Il existe donc un borélien  $B_*$  de  $\Sigma_A$  tel que  $\mu'(B_*) > 0$  et  $\mu(B_*) = 0$ . Rappelons que l'ensemble  $\mathcal{E}$  des unions finies disjointes de cylindres est une algèbre de Boole engendrant le  $\sigma$ -algèbre produit de  $\Sigma$ .

Le lemme 5.2 appliqué à la mesure  $\frac{1}{2}(\mu' + \mu)$ , permet de trouver une suite  $(B_p)_{p \in \mathbb{N}}$  d'unions finies disjointes de cylindres de  $\Sigma_A$  telles que  $\lim_{p \rightarrow \infty} \mu'(B_p) = \mu'(B_*) \neq 0$  et  $\lim_{p \rightarrow \infty} \mu(B_p) = 0$ .

Notons  $\xi_0$  la partition mesurable de  $\Sigma_A$  définie par

$$\xi_0 = \{\Sigma_A \cap [\alpha]_0 : \alpha \in \mathcal{A}, \Sigma_A \cap [\alpha]_0 \neq \emptyset\}$$

et  $\xi_n$  la partition  $\xi_n = \bigvee_{i=-n}^n \sigma_A^{-i} \xi_0$ . Comme  $\xi_0$  est une partition génératrice pour la transformation inversible  $\sigma_A$ , d'après le théorème de Kolmogorov-Sinai (proposition 5.3 (8)) et la proposition 5.3 (1) dans le cas inversible, nous avons

$$h_{\mu'}(\sigma_A) = h_{\mu'}(\sigma_A, \xi_0) = \inf_{n \geq 1} \frac{1}{2n+1} H_{\mu'}(\xi_n).$$

Nous avons donc, pour tout  $n \geq 1$ , l'inégalité  $H_{\mu'}(\xi_n) - (2n+1) \ln \lambda \geq 0$ .

**Lemme 11.25.** Soient  $x_1, \dots, x_k$  des réels strictement positifs tels que  $\sum_{i=1}^k x_i = a$ . Alors  $-\sum_{i=1}^k x_i \log x_i \leq a \log \frac{k}{a}$ .

**Démonstration.** Par la concavité de la fonction  $\varphi : t \mapsto -t \log t$  sur  $[0, 1]$ , nous avons

$$\frac{1}{k} \sum_{i=1}^k -x_i \log x_i \leq -\left(\frac{1}{k} \sum_{i=1}^k x_i\right) \log \left(\frac{1}{k} \sum_{i=1}^k x_i\right) = \frac{a}{k} \log \frac{k}{a}. \quad \square$$

**Lemme 11.26.** Il existe  $c > 0$  tel que, pour tout cylindre  $C = [\alpha_m, \dots, \alpha_n]_m$  non vide de  $\Sigma_A$ , nous ayons  $\mu(C) \geq c \lambda^{-(n-m)}$ .

**Démonstration.** Par la définition (138) de la mesure de Parry et le lemme 6.3 appliqué avec  $\pi = \pi^\bullet$  et  $\nu = \nu^{p_i^\bullet}$ , par la remarque 11.23 et la formule (137), et puisque  $f_{\alpha_m} \prod_{i=m}^{n-1} \frac{\alpha_{i+1}}{\alpha_i} = f_{\alpha_n}$ , nous avons

$$\mu(C) = \nu_{\alpha_m}^{\pi^\bullet} \prod_{i=m}^{n-1} \pi_{\alpha_i, \alpha_{i+1}}^\bullet = \left( \prod_{i=m}^{n-1} A_{\alpha_i, \alpha_{i+1}} \right) v_{\alpha_m} f_{\alpha_n} \lambda^{-(n-m)}.$$

Il suffit donc de prendre  $c = \min_{\alpha, \beta \in \mathcal{A}} v_\alpha f_\beta > 0$ . □

**Fin de la démonstration de l'assertion b) du théorème 11.24.** Fixons  $p$  et choisissons  $n$  de sorte que  $B_p$  soit une réunion finie de parties de la partition  $\xi_n$ . Nous avons alors, en utilisant les deux lemmes précédents,

$$\begin{aligned} 0 &\leq H_{\mu'}(\xi_n) - (2n+1) \log \lambda \\ &= - \sum_{C \in \xi_n, C \subset B_p} \mu'(C) \ln \mu'(C) - \sum_{C \in \xi_n, C \subset {}^c B_p} \mu'(C) \ln \mu'(C) - (2n+1) \ln \lambda \\ &\leq \mu'(B_p) \ln \frac{\text{Card} \{C \in \xi_n : C \subset B_p\}}{\mu'(B_p)} + \mu'({}^c B_p) \ln \frac{\text{Card} \{C \in \xi_n : C \subset {}^c B_p\}}{\mu'({}^c B_p)} - (2n+1) \ln \lambda \\ &\leq \mu'(B_p) \ln \frac{\mu(B_p) \lambda^{2n}}{c \mu'(B_p)} + \mu'({}^c B_p) \ln \frac{\mu({}^c B_p) \lambda^{2n}}{c \mu'({}^c B_p)} - (2n+1) \ln \lambda \\ &= \mu'(B_p) \ln \frac{\mu(B_p)}{\mu'(B_p)} + \mu'({}^c B_p) \ln \frac{\mu({}^c B_p)}{\mu'({}^c B_p)} - \ln c - \ln \lambda. \end{aligned}$$

Ces inégalités sont vraies pour tout  $p$ . Mais le dernier terme tend vers  $-\infty$  quand  $p$  tend vers l'infini, car  $\lim_{p \rightarrow +\infty} \mu(B_p) = 0$ ,  $\lim_{p \rightarrow +\infty} \mu'({}^c B_p) > 0$  et  $\mu'({}^c B_p) \ln \frac{\mu({}^c B_p)}{\mu'({}^c B_p)}$  est borné supérieurement par  $\frac{1}{e}$ . Ceci donne la contradiction cherchée.

**c)** Comme les fonctions localement constantes sont denses (pour la norme uniforme) dans les fonctions continues sur l'espace de Cantor  $\mathcal{A}^{\mathbb{Z}}$ , il suffit de montrer que, pour tout cylindre  $C = [\alpha_m, \dots, \alpha_n]_m$  de  $\mathcal{A}^{\mathbb{Z}}$ , nous avons  $\lim_{N \rightarrow \infty} \mu_N(C) = \mu(C)$ .

Puisque de tout sommet du graphe orienté  $\mathcal{G}_A$  part et arrive au moins une arête, il est possible de compléter la suite finie  $(\alpha_m, \dots, \alpha_n)$  en un élément de  $\Sigma_A$  si et seulement si  $\prod_{i=m}^{n-1} A_{\alpha_i, \alpha_{i+1}} = 1$ . Dans ce cas et si  $N > n - m$ , cette suite finie se complète en une suite

$$(\alpha'_m, \dots, \alpha'_n, \alpha'_{n+1}, \dots, \alpha'_{N+m-1})$$

de longueur  $N$  qui elle-même se complète en un élément de  $P_N \cap C$ , si et seulement si  $\alpha'_m = \alpha_m, \dots, \alpha'_n = \alpha_n, \alpha'_{N+m-1} = \alpha_m$  et  $(\alpha'_n, \alpha'_{n+1}, \dots, \alpha'_{N+m-1})$  appartient à

$$\begin{aligned} \{(\beta_n, \dots, \beta_{N+m-1}) \in \mathcal{A}^{N-n+m} : \beta_0 = \alpha_n, \beta_n = \alpha_m \text{ et} \\ [\beta_n]_0 \cap \sigma_A^{-1}([\beta_{n+1}]_0) \cap \dots \cap \sigma_A^{-N+n-m+1}([\beta_{N+m-1}]_0) \neq \emptyset\}. \end{aligned}$$

D'après le lemme 11.21, nous avons donc les égalités  $\text{Card } P_N = \text{tr } A^N$  et, pour  $N > n-m$ ,

$$\mu_N(C) = \frac{\text{Card } (P_N \cap C)}{\text{Card } P_N} = \frac{1}{\text{tr } A^N} \left( \prod_{i=m}^{n-1} A_{\alpha_i \alpha_{i+1}} \right) (A^{N-n+m})_{\alpha_n \alpha_m} .$$

Soient  $f = (f_\alpha)_{\alpha \in \mathcal{A}}$  et  $v = (v_\alpha)_{\alpha \in \mathcal{A}}$  définis juste avant la formule (135). Notons<sup>250</sup>  $\tau$  le projecteur sur la droite  $\mathbb{R}f$  parallèlement à l'hyperplan  $H$  noyau de la forme linéaire duale à  $v$ , défini par

$$H = \{ (x_\alpha)_{\alpha \in \mathcal{A}} \in \mathbb{R}^{\mathcal{A}} : \sum_{\alpha \in \mathcal{A}} v_\alpha x_\alpha = 0 \} .$$

D'après le théorème de Perron-Frobenius, nous avons  $\lim_{N \rightarrow \infty} \frac{1}{\lambda^N} A^N = \tau$ . En effet, nous avons  $\lim_{N \rightarrow \infty} \frac{1}{\lambda^N} A^N f = f$ . Puisque  $v$  est fixé par  ${}^t A$ , l'hyperplan  $H$  est invariant par  $A$  : si  $x \in H$ , avec  $\langle \cdot, \cdot \rangle$  le produit scalaire euclidien canonique (rendant la base canonique orthonormée) de  $\mathbb{R}^{\mathcal{A}}$ , nous avons  $\langle Ax, v \rangle = \langle x, {}^t Av \rangle = \langle x, v \rangle = 0$ . Comme les valeurs propres de  $A$  distinctes de  $\lambda$  sont de module strictement inférieur à  $\lambda$ , et comme  $\lambda$  est de multiplicité 1, la restriction de  $A$  à  $H$  est donc de rayon spectral strictement inférieur à  $\lambda$ . Donc  $\lim_{N \rightarrow \infty} \frac{1}{\lambda^N} A^N x = 0$  pour tout  $x \in H$ .

Nous avons  $\tau = (f_\alpha v_\beta)_{\alpha, \beta \in \mathcal{A}}$  car si  $\tau' = (f_\alpha v_\beta)_{\alpha, \beta \in \mathcal{A}}$ , alors  $\tau' f = f$  par la formule (135), et pour tout  $x \in H$ , nous avons  $\tau' x = 0$ , donc  $\tau' = \tau$ .

En particulier, nous avons que  $\lim_{N \rightarrow \infty} \frac{1}{\lambda^N} \text{tr } A^N = \text{tr } \tau = 1$  par la formule (135) ainsi que  $\lim_{N \rightarrow \infty} \frac{1}{\lambda^N} (A^N)_{\alpha, \beta} = f_\alpha v_\beta$ . D'où par la formule (137), par une simplification télescopique du troisième terme et par le lemme 6.3, nous avons

$$\lim_{N \rightarrow \infty} \mu_N(C) = \left( \prod_{i=m}^{n-1} A_{\alpha_i \alpha_{i+1}} \right) \frac{v_{\alpha_m} f_{\alpha_n}}{\lambda^{n-m}} = f_{\alpha_m} v_{\alpha_m} \prod_{i=m}^{n-1} \pi_{\alpha_i \alpha_{i+1}} = \mu(C) .$$

Ceci démontre l'assertion c). □

## 11.4 Entropie topologique des automorphismes linéaires du tore

Cette partie est consacrée à l'étude de l'entropie topologique, de l'existence et l'unicité d'une mesure d'entropie maximale et des conséquences sur l'équidistribution des points périodiques pour les systèmes dynamiques (uniformément) hyperboliques. Nous décrirons essentiellement la situation du tore  $\mathbb{T}^N$  muni d'un automorphisme linéaire hyperbolique  $\phi_M$ , déjà moultement étudié dans ces notes. Il est de toute façon primordial de se familiariser avec cet exemple pour comprendre le cas général des systèmes dynamiques hyperboliques.

Nous procéderons par codage (au sens de l'introduction de la partie 11.3) du système dynamique  $(\mathbb{T}^N, \phi_M)$  par un sous-décalage de type fini  $(\Sigma_A, \sigma_A)$ , défini à l'aide de "partitions de Markov" construites à partir des variétés stables et instables de  $(\mathbb{T}^N, \phi_M)$ . La démonstration de l'existence de telles partitions représente l'essentiel du travail.

Nous utiliserons alors l'analyse des sous-décalages de type fini effectuée dans la partie 11.3 pour montrer que la mesure de Haar sur  $\mathbb{T}^N$  est l'unique mesure d'entropie maximale

---

250. en identifiant les endomorphismes linéaires de  $\mathbb{R}^{\mathcal{A}}$  avec leur matrice dans la base canonique  $(e_\alpha)_{\alpha \in \mathcal{A}}$  où  $e_\alpha(\beta) = \begin{cases} 1 & \text{si } \alpha = \beta \\ 0 & \text{sinon} \end{cases}$  de  $\mathbb{R}^{\mathcal{A}}$ , et  $\mathbb{R}^{\mathcal{A}}$  avec son dual par l'isomorphisme envoyant la base canonique de  $\mathbb{R}^{\mathcal{A}}$  sur sa base duale

pour  $\phi_M$ , pour calculer l'entropie topologique de  $(\mathbb{T}^N, \phi_M)$ , et pour montrer l'équidistribution des points périodiques de  $\phi_M$  vers la mesure de Haar sur  $\mathbb{T}^N$ .

#### 11.4.1 Partition de Markov du tore

Dans toute la suite de cette partie, nous reprenons les notations de la partie 10.3 sur la stabilité structurelle des automorphismes linéaires hyperboliques du tore. En particulier, nous notons

- $N$  un élément de  $\mathbb{N} - \{0\}$  et  $M$  une matrice  $N \times N$  à coefficients entiers, hyperbolique (c'est-à-dire sans valeur propre de module 1), que nous supposons de plus de déterminant égal à  $\pm 1$  (pour que  $\phi_M$  soit inversible), par exemple  $M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ ;
- $\mathbb{R}^N = E^s \oplus E^u$  la décomposition en somme directe de  $\mathbb{R}^N$  en sous-espaces vectoriels stables et instables pour  $M$ ;
- $\|\cdot\|$  une norme sur  $\mathbb{R}^N$  adaptée à  $M$  (c'est-à-dire telle que, si nous notons encore  $\|\cdot\|$  la norme d'opérateur associée, alors  $\|M|_{E^s}\| < 1$ ,  $\|M^{-1}|_{E^u}\| < 1$  et, pour tous les  $x_s$  dans  $E^s$  et  $x_u$  dans  $E^u$ , nous ayons  $\|x_s + x_u\| = \max\{\|x_s\|, \|x_u\|\}$ )<sup>251</sup>. Nous noterons  $\overline{B}(v, r)$  la boule fermée de centre  $v \in \mathbb{R}^N$  et de rayon  $r > 0$  pour la norme  $\|\cdot\|$ ;
- $\text{ch}(M)$  la constante d'hyperbolicité de  $M$

$$\text{ch}(M) = \max\{\|M|_{E^s}\|, \|M^{-1}|_{E^u}\|\} < 1;$$

- $\pi : \mathbb{R}^N \rightarrow \mathbb{T}^N$  la projection canonique;
- $\phi = \phi_M$  l'automorphisme linéaire hyperbolique du tore  $\mathbb{T}^N$  défini par  $\phi_M \circ \pi = \pi \circ M$ ;
- et
- $d$  la distance quotient sur  $\mathbb{T}^N$

$$d(x, y) = \inf\{\|v - w\| : \pi(v) = x, \pi(w) = y\}.$$

Fixons  $\epsilon_0 > 0$  tel que l'application  $\pi$  soit isométrique (et en particulier injective) sur la boule  $\overline{B}(0, \epsilon_1) = \{v \in \mathbb{R}^N : \|v\| \leq \epsilon_1\}$  dans  $\mathbb{R}^N$  de rayon

$$\epsilon_1 = 2\epsilon_0 \max\{1, \|M\|, \|M^{-1}\|\}. \quad (140)$$

Elle est alors isométrique sur toute boule fermée de rayon  $\epsilon_1$ .

Pour  $\epsilon < \epsilon_0$  et  $x$  dans  $\mathbb{T}^N$ , notons, comme cas particulier de la situation générale décrite dans la partie 10.8,

$$W_\epsilon^s(x) = \{y \in \mathbb{T}^N : \forall n \geq 0, d(\phi^n(x), \phi^n(y)) \leq \epsilon\}$$

la *variété stable locale* en  $x$  et

$$W_\epsilon^u(x) = \{y \in \mathbb{T}^N : \forall n \geq 0, d(\phi^{-n}(x), \phi^{-n}(y)) \leq \epsilon\}$$

la *variété instable locale* en  $x$ .

Rassemblons quelques propriétés élémentaires de ces sous-espaces.

---

251. Attention, il ne s'agit pas d'une norme euclidienne.

**Lemme 11.27.** Soient  $\epsilon < \epsilon_0$  et  $x, y$  dans  $\mathbb{T}^N$ .

(1) Nous avons  $\phi(W_\epsilon^s(x)) \subset W_\epsilon^s(\phi(x))$  et  $W_\epsilon^u(\phi(x)) \subset \phi(W_\epsilon^u(x))$ .

(2) Le point  $y$  appartient à  $W_\epsilon^s(x)$  si et seulement si pour tout  $n \geq 0$ , nous avons

$$d(\phi^n(x), \phi^n(y)) \leq \epsilon \operatorname{ch}(M)^n .$$

(3) Le point  $y$  appartient à  $W_\epsilon^u(x)$  si et seulement si pour tout  $n \geq 0$ , nous avons

$$d(\phi^{-n}(x), \phi^{-n}(y)) \leq \epsilon \operatorname{ch}(M)^n .$$

(4) Si  $d(x, y) \leq \epsilon$ , alors l'intersection  $W_\epsilon^s(x) \cap W_\epsilon^u(y)$  est réduite à un point, noté  $[x, y]$ . L'application de  $\{(x, y) \in \mathbb{T}^N \times \mathbb{T}^N : d(x, y) \leq \epsilon\}$  dans  $\mathbb{T}^N$  définie par  $(x, y) \mapsto [x, y]$  est continue, et appelée le produit local associé à  $\phi_M$ .

(5) Si  $\mathcal{P}(v) = x$ , la variété stable en  $x$  est

$$W^s(x) = \{y \in \mathbb{T}^N : \lim_{n \rightarrow +\infty} d(\phi^n(x), \phi^n(y)) = 0\} = \bigcup_{n \in \mathbb{N}} \phi^{-n}(W_\epsilon^s(\phi^n(x))) = \mathcal{P}(v + E^s)$$

et la variété instable en  $x$  est

$$W^u(x) = \{y \in \mathbb{T}^N : \lim_{n \rightarrow +\infty} d(\phi^{-n}(x), \phi^{-n}(y)) = 0\} = \bigcup_{n \in \mathbb{N}} \phi^n(W_\epsilon^u(\phi^{-n}(x))) = \mathcal{P}(v + E^u) .$$

Notons que les variétés stables et instables  $W^s(x)$  et  $W^u(x)$  sont des sous-variétés immergées comme dicté par la théorie générale (voir la partie 10.8), mais ne sont en général pas des sous-variétés plongées : elles peuvent être denses dans le tore  $\mathbb{T}^N$ , comme c'est par exemple le cas pour l'application du chat d'Arnold  $\phi_M$  donnée par  $M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ , où les droites stables et instables de  $M$  dans  $\mathbb{R}^2$  sont de pente irrationnelle (voir l'exercice E.62).

**Démonstration.** Il est possible aussi de donner une description explicite des variétés stables et instables locales. Fixons  $\epsilon \leq \epsilon_0$ . Pour tous les points  $x, y$  de  $\mathbb{T}^N$ , notons  $\tilde{x}, \tilde{y}$  des relevés dans  $\mathbb{R}^N$  de  $x, y$ , de sorte que si  $d(x, y) \leq \epsilon$ , alors  $\|\tilde{x} - \tilde{y}\| \leq \epsilon$ . Nous allons montrer les égalités

$$W_\epsilon^s(x) = \mathcal{P}(\overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s)) \quad \text{et} \quad W_\epsilon^u(x) = \mathcal{P}(\overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^u)) , \quad (141)$$

dont découle en particulier l'assertion (5).

(1) Ceci est immédiat par les définitions de  $W_\epsilon^s(x)$  et  $W_\epsilon^u(x)$ .

(2) Soit  $v \in B(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s)$ . Par la définition de la distance quotient  $d$ , par la linéarité de  $M$ , puisque  $\tilde{x} - \tilde{v} \in E^s$  et puisque  $M|_{E^s}$  est  $\operatorname{ch}(M)$ -lipschitzienne sur  $E^s$ , pour tout  $n \in \mathbb{N}$ , nous avons

$$d(\phi^n(x), \phi^n(\mathcal{P}(v))) \leq \|M^n \tilde{x} - M^n \tilde{v}\| \leq \operatorname{ch}(M)^n \|\tilde{x} - \tilde{v}\| \leq \epsilon \operatorname{ch}(M)^n \leq \epsilon .$$

Donc

$$\mathcal{P}(\overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s)) \subset \{y \in \mathbb{T}^N : \forall n \in \mathbb{N}, d(\phi^n(x), \phi^n(y)) \leq \epsilon \operatorname{ch}(M)^n\} \subset W_\epsilon^s(x) . \quad (142)$$

Réciproquement, si  $y \in W_\epsilon^s(x)$ , alors en particulier  $d(x, y) \leq \epsilon$ . Comme dit ci-dessus, nous choisissons des relevés  $\tilde{x}, \tilde{y}$  de  $x, y$ , de sorte que  $\|\tilde{x} - \tilde{y}\| \leq \epsilon$ . Notons  $\tilde{y} - \tilde{x} = v_s + v_u$  avec  $v_s \in E^s$  et  $v_u \in E^u$ . Supposons par l'absurde que  $v_u \neq 0$ . Puisque  $M|_{E^u}$  est dilatante, il existe un entier  $n \in \mathbb{N}$  tel que  $M^n v_u \in \overline{B}(0, \epsilon)$  et  $M^{n+1} v_u \notin \overline{B}(0, \epsilon)$ . Notons que  $M^{n+1} v_u \in \overline{B}(0, \epsilon \|M\|)$  et que par la définition (140) de  $\epsilon_1$ , l'application  $\rho$  est une isométrie sur  $B(M^{n+1} \tilde{x}, \epsilon_0 \|M\|)$ . Donc, puisque la norme est adaptée, nous avons

$$d(\phi^{n+1}(y), \phi^{n+1}(x)) = \|M^{n+1} \tilde{y} - M^{n+1} \tilde{x}\| = \|M^{n+1} v_s + M^{n+1} v_u\| \geq \|M^{n+1} v_u\| > \epsilon,$$

ce qui contredit le fait que  $y \in W_\epsilon^s(x)$ . Donc  $\tilde{y} - \tilde{x} = v_s \in E^s$  et  $\tilde{y} \in \overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s)$ . Par conséquent  $y \in \rho(\overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s))$ . Ceci montre que la chaîne d'inclusions dans la formule (142) est en fait une chaîne d'égalités. L'assertion (2) en découle, ainsi que l'égalité de gauche dans la formule (141).

(3) L'assertion (3), ainsi que l'égalité de droite dans la formule (141), découlent de la démonstration précédente, en remplaçant  $\phi$  par  $\phi^{-1} = \phi_{M^{-1}}$ , et en remarquant que  $\text{ch}(M) = \text{ch}(M^{-1})$  et que la définition (140) de  $\epsilon_1$  est symétrique en  $M$  et  $M^{-1}$ .

(4) Comme vu dans les préliminaires, soient  $\tilde{x}, \tilde{y}$  des relevés de  $x, y$  tels que  $\|\tilde{x} - \tilde{y}\| \leq \epsilon$ . Les sous-espaces affines  $\tilde{x} + E^s$  et  $\tilde{y} + E^u$ , dont les directions sont supplémentaires, se coupent en un et un seul point  $\tilde{z}$ . De plus, le point  $\tilde{z}$  dépend continuellement de  $\tilde{x}, \tilde{y}$ . Puisque la norme est adaptée, puisque  $\tilde{x} - \tilde{z} \in E^s$  et  $\tilde{z} - \tilde{y} \in E^u$ , nous avons

$$\max \{ \|\tilde{x} - \tilde{z}\|, \|\tilde{z} - \tilde{y}\| \} = \|(\tilde{x} - \tilde{z}) + (\tilde{z} - \tilde{y})\| = \|\tilde{x} - \tilde{y}\| \leq \epsilon$$

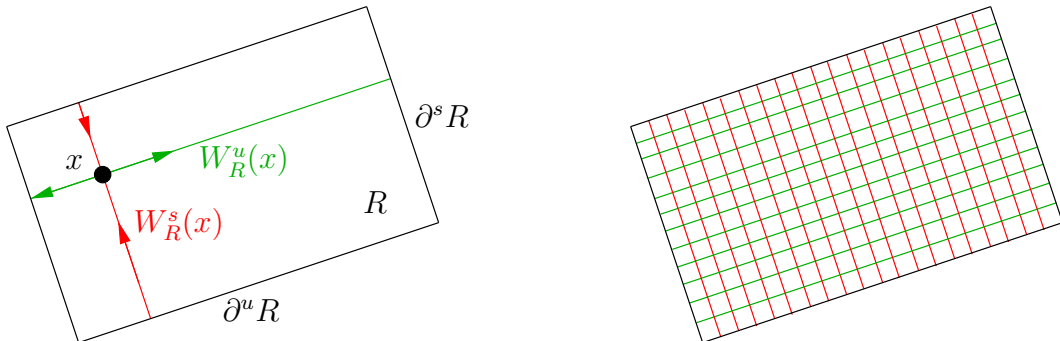
Par les formules (141), nous avons donc  $z = \rho(\tilde{z}) \in W_\epsilon^s(x) \cap W_\epsilon^u(y)$ . Si  $z'$  est un autre point de  $W_\epsilon^s(x) \cap W_\epsilon^u(y)$ , alors  $z'$  admet un unique relevé  $\tilde{z}'$  appartenant à  $\overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s)$  et il existe un relevé  $\tilde{y}'$  de  $y$  tel que  $\tilde{z}' \in \overline{B}(\tilde{y}', \epsilon) \cap (\tilde{y}' + E^u)$ . Donc comme ci-dessus

$$\|\tilde{y}' - \tilde{x}\| = \|(\tilde{y}' - \tilde{z}') + (\tilde{z}' - \tilde{x})\| = \max \{ \|\tilde{x} - \tilde{z}'\|, \|\tilde{z}' - \tilde{y}'\| \} \leq \epsilon.$$

Donc les points  $\tilde{y}'$  et  $\tilde{y}$  sont tous deux à distance au plus  $\epsilon < \epsilon_0$  de  $\tilde{x}$ , et s'envoient par  $\rho$  sur le même point  $y$ . Puisque  $\rho$  est injective sur la boule  $\overline{B}(\tilde{x}, \epsilon)$  par la définition (140) de  $\epsilon_1$ , ceci implique que  $\tilde{y}' = \tilde{y}$  et  $z' = z$ .  $\square$

Nous appellerons *rectangle* de  $\mathbb{T}^N$  pour  $\phi = \phi_M$  toute partie  $R$  non vide de diamètre au plus  $\epsilon_0$  telle que si  $x, y \in R$ , alors  $[x, y] \in R$ .

Autrement dit, dans une identification locale de  $\mathbb{T}^N$  avec  $E = E^s \times E^u$ , nous avons  $R = R^s \times R^u$  avec  $\text{diam}(R^s) \leq \epsilon_0$  et  $\text{diam}(R^u) \leq \epsilon_0$ . Les parties  $R^s$  et  $R^u$  ne sont pas forcément des boules de  $E^s$  et  $E^u$ . La frontière de  $R$  s'écrit alors  $\partial R = \partial^s R \cup \partial^u R$  où  $\partial^s R = R^s \times \partial R^u$  et  $\partial^u R = \partial R^s \times R^u$ .

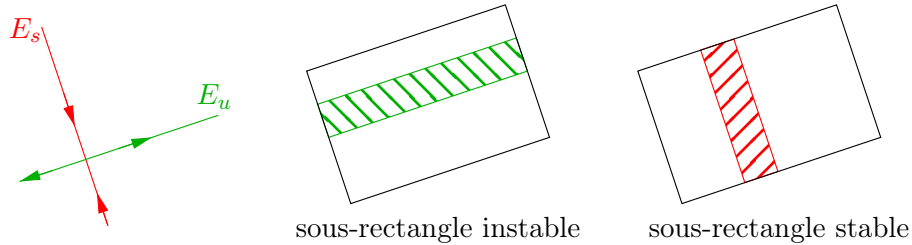




Le rectangle  $R$  est dit *propre* s'il est l'adhérence de son intérieur. Pour  $x$  dans  $R$ , notons

$$W_R^s(x) = W_{\epsilon_0}^s(x) \cap R \quad \text{et} \quad W_R^u(x) = W_{\epsilon_0}^u(x) \cap R,$$

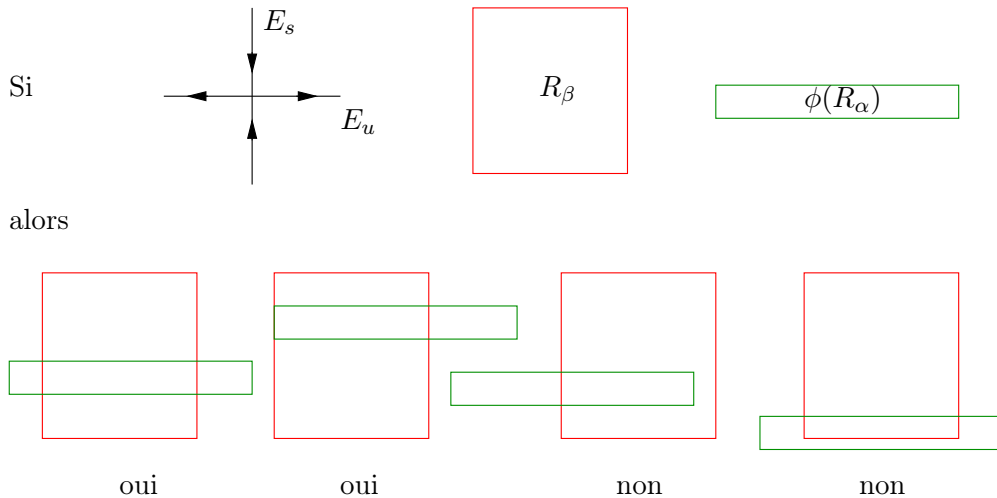
que nous appellerons les feuilles stables et instables dans  $R$  de  $x$ . Une partie  $S$  de  $R$  est appelée un *sous-rectangle stable* (resp. *instable*) si pour tout  $x$  dans  $S$ , la partie  $W_R^s(x)$  (resp.  $W_R^u(x)$ ) est contenue dans  $S$ .



Une *partition de Markov*<sup>252</sup>  $\mathcal{R}$  de  $\mathbb{T}^N$  pour  $\phi$  est un recouvrement fini  $\mathcal{R} = (R_\alpha)_{\alpha \in \mathcal{A}}$  de  $\mathbb{T}^N$  par des rectangles propres de  $\mathbb{T}^N$  pour  $\phi$ , tel que, pour tous les  $\alpha, \beta$  dans  $\mathcal{A}$ , nous ayons

- a)  $\overset{\circ}{R}_\alpha \cap \overset{\circ}{R}_\beta = \emptyset$  si  $\alpha \neq \beta$ ,
- b) si  $x \in \overset{\circ}{R}_\alpha$  et  $\phi(x) \in \overset{\circ}{R}_\beta$ , alors  $\phi(W_{R_\alpha}^s(x)) \subset W_{R_\beta}^s(\phi(x))$  et  $W_{R_\beta}^u(\phi(x)) \subset \phi(W_{R_\alpha}^u(x))$ .

Voici un dessin expliquant graphiquement la condition b).



Nous renvoyons par exemple à [BenP, §10] pour une démonstration du résultat suivant.

**Théorème 11.28.** *Pour tout  $\epsilon > 0$  et tout automorphisme linéaire hyperbolique  $\phi_M$  de  $\mathbb{T}^N$ , il existe une partition de Markov  $\mathcal{R}$  pour  $\phi_M$  par des rectangles de diamètre au plus  $\epsilon$ .  $\square$*

252. La terminologie “partition de Markov” est classique, bien que  $\mathcal{R}$  ne soit jamais une partition de  $\mathbb{T}^N$  au sens strict, et que, étant un objet de dynamique topologique, elle n’est pas initialement reliée à des chaînes de Markov.

### 11.4.2 Codage des automorphismes linéaires hyperboliques du tore

À chaque partition de Markov, nous associons sa *matrice d'incidence*  $A = (A_{\alpha\beta})_{\alpha,\beta \in \mathcal{A}}$  où, pour tous les  $\alpha, \beta \in \mathcal{A}$ , nous avons

$$A_{\alpha\beta} = \begin{cases} 1 & \text{si } \phi(R_\alpha) \cap R_\beta \neq \emptyset \\ 0 & \text{sinon.} \end{cases}$$

**Théorème 11.29.** Soient  $\mathcal{R} = (R_\alpha)_{\alpha \in \mathcal{A}}$  une partition de Markov pour  $\phi = \phi_M$  par des rectangles de diamètre au plus  $\epsilon$  avec  $\epsilon > 0$  assez petit,  $A$  sa matrice d'incidence et  $(\Sigma_A, \sigma_A)$  le sous-décalage de type fini associé.

- Pour tout  $\omega$  dans  $\Sigma_A$ , l'intersection  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(R_{\omega_n})$  est réduite à un point, que nous noterons  $h(\omega)$ .
- L'application  $h : \Sigma_A \rightarrow \mathbb{T}^N$  ainsi définie est continue, surjective et vérifie la propriété de semi-conjugaison  $\phi \circ h = h \circ \sigma_A$  entre  $\sigma_A$  et  $\phi$ .
- Pour toute mesure de probabilité  $\sigma_A$ -invariante et ergodique  $\mu$ , de support  $\Sigma_A$ , nous avons

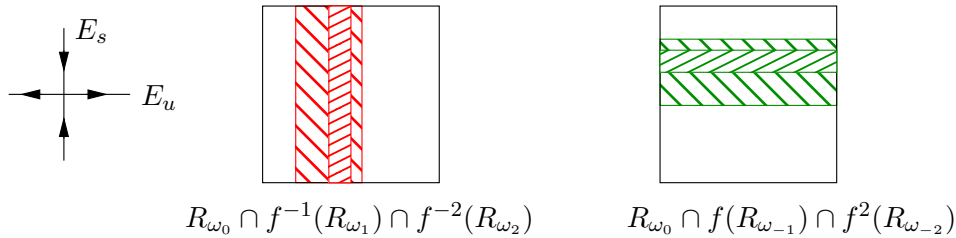
$$\mu(\{\omega \in \Sigma_A : \text{Card } h^{-1}(h(\omega)) > 1\}) = 0 .$$

Ce théorème (avec l'existence de partitions de Markov vue en théorème 11.28) dit en particulier que pour tout automorphisme linéaire hyperbolique du tore, il existe un sous-décalage de type fini qui lui est (topologiquement) semi-conjugué. De plus, quitte à enlever un ensemble négligeable (pour toute mesure raisonnable), cette semi-conjugaison est en fait une conjugaison.

**Démonstration.** Nous noterons

$$\overset{\circ}{\mathcal{R}} = \bigcup_{\alpha \in \mathcal{A}} \overset{\circ}{R}_\alpha, \quad \partial \mathcal{R} = \bigcup_{\alpha \in \mathcal{A}} \partial R_\alpha, \quad \partial^s \mathcal{R} = \bigcup_{\alpha \in \mathcal{A}} \partial^s R_\alpha, \quad \partial^u \mathcal{R} = \bigcup_{\alpha \in \mathcal{A}} \partial^u R_\alpha .$$

a) Soit  $\omega = (\omega_n)_{n \in \mathbb{Z}}$  un élément de  $\Sigma_A$ . Par la définition de  $\Sigma_A$  et la propriété b) des partitions de Markov, la partie  $R_{\omega_0} \cap \phi^{-1}(R_{\omega_1})$  est un sous-rectangle stable non vide de  $R_{\omega_0}$ . Par récurrence, l'intersection  $\bigcap_{n=0}^{n_0} \phi^{-n}(R_{\omega_n})$  est un sous-rectangle stable non vide pour tout  $n_0 \in \mathbb{N}$ . Par compacité, l'intersection  $\bigcap_{n \geq 0} \phi^{-n}(R_{\omega_n})$  est aussi un sous-rectangle stable non vide.



Nous montrons de même que  $\bigcap_{n \leq 0} \phi^{-n}(R_{\omega_n})$  est un sous-rectangle instable non vide. Donc l'intersection  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(R_{\omega_n})$  est non vide. Deux points de cette intersection ont des orbites qui se suivent à  $\epsilon$  près (puisque les diamètres des rectangles sont au plus  $\epsilon$ ), et sont donc égaux (par le lemme de pistage 10.20) si  $\epsilon$  est assez petit.

b) Par compacité, il résulte de a) que pour tout  $\omega = (\omega_n)_{n \in \mathbb{Z}}$ , nous avons

$$\lim_{n_0 \rightarrow +\infty} \text{diam} \left( \bigcap_{n=-n_0}^{n_0} \phi^{-n}(R_{\omega_n}) \right) = 0 .$$

Ceci prouve la continuité de  $h$  par la définition de la distance sur  $\Sigma_A$  induite de celle sur  $\Sigma = \mathbb{A}^{\mathbb{Z}}$  (voir la formule (4)). Par construction, nous avons  $\phi \circ h = h \circ \sigma_A$ . Par construction, tout point de l'intersection  $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(\overset{\circ}{\mathcal{R}})$  appartient à l'image  $h(\Sigma_A)$  de  $h$ . Par le théorème de Baire, cette intersection est dense dans  $\mathbb{T}^N$ . Comme  $h(\Sigma_A)$  est compact, l'application  $h$  est surjective.

c) La mesure de probabilité image  $\mu' = h_*\mu$  est donc invariante par  $\phi$ , ergodique et de support  $\mathbb{T}^N$ . Soit  $Z = \{x \in \mathbb{T}^N : \text{Card } h^{-1}(x) > 1\}$ . Nous voulons montrer que  $\mu'(Z) = 0$ . Par construction, l'ensemble  $Z$  est contenu dans  $\bigcup_{n \in \mathbb{Z}} \phi^n(\partial \mathcal{R})$ . Écrivons  $\partial \mathcal{R} = \partial^u \mathcal{R} \cup \partial^s \mathcal{R}$ . La propriété b) des partitions de Markov implique que  $\phi(\partial^s \mathcal{R}) \subset \partial^s \mathcal{R}$ . Comme  $\mu'$  est invariante par  $\phi$ , nous avons  $\mu'(\partial^s \mathcal{R}) = \mu'(\phi^n(\partial^s \mathcal{R})) = \mu'(\bigcap_{n \geq 0} \phi^n(\partial^s \mathcal{R}))$ . Or la partie fermée  $F = \bigcap_{n \geq 0} \phi^n(\partial^s \mathcal{R})$  vérifie  $\phi^{-1}(F) = F$  et n'est pas de mesure 1, car le support de  $\mu'$  est égal à  $\mathbb{T}^N$ . Par ergodicité, elle est donc de mesure nulle, et donc  $\mu'(\partial^s \mathcal{R}) = 0$ . De même, nous avons  $\mu'(\partial^u \mathcal{R}) = 0$ , donc  $\mu'(\partial \mathcal{R}) = 0$  et  $\mu'(Z) = 0$  par union dénombrable d'ensembles de mesure nulle.  $\square$

**Remarque.** (1) Sous les hypothèses du théorème, nous pouvons montrer que, pour tout  $x$  dans  $\mathbb{T}^N$ , nous avons

$$\text{Card } h^{-1}(x) \leq (\text{Card } \mathcal{A})^2 .$$

(2) Comme  $\phi = \phi_M$  est topologiquement mélangeante (par le théorème 4.10 et l'exercice E.50), le sous-décalage  $\sigma_A$  est aussi topologiquement mélangeant et la matrice  $A$  est irréductible (voir l'exercice E.38). Ceci permet, pour toute matrice stochastique  $\Pi = (\pi_{\alpha\beta})_{\alpha, \beta \in \mathcal{A}}$  telle que  $\pi_{\alpha\beta} = 0$  si et seulement si  $A_{\alpha\beta} = 0$ , de construire une mesure de probabilité invariante par  $\sigma_A$  et ergodique  $\mathbb{P}_\Pi$  de support  $\Sigma_A$ . Par l'exercice E.39, en prenant l'image dans  $\mathbb{T}^N$  de ces mesures de probabilité, nous construisons une famille à plusieurs paramètres réels (donc une famille non dénombrable) de mesures de probabilité images  $h_*\mathbb{P}_\Pi$  sur  $\mathbb{T}^N$  qui sont  $\phi_M$ -invariantes, ergodiques et deux à deux étrangères.

### 11.4.3 Mesure d'entropie maximale des automorphismes hyperboliques du tore

**Proposition 11.30.** Soient  $\phi_M$  un automorphisme linéaire hyperbolique du tore  $\mathbb{T}^N$  et  $\mu$  la mesure de probabilité de Haar sur  $\mathbb{T}^N$ . Notons  $\lambda_+ = |\det(M|_{E^u})|$ . Alors

$$h_{\text{top}}(\phi_M) = h_\mu(\phi_M) = \ln \lambda_+ .$$

De plus,  $\mu$  est l'unique mesure d'entropie maximale de  $\phi_M$  sur  $\mathbb{T}^N$ .

**Remarque.** Nous avons  $\log \lambda_+ = \sum \log |\lambda_i|$ , où les  $\lambda_i$  sont les valeurs propres complexes de  $M$  telles que  $|\lambda_i| > 1$ .

**Démonstration.** Soit  $\mathcal{R} = (R_\alpha)_{\alpha \in \mathcal{A}}$  une partition de Markov pour  $\phi_M$ . Choisissons  $\xi_0 = \{R'_\alpha : \alpha \in \mathcal{A}\}$  une (vraie) partition mesurable de  $\mathbb{T}^N$  telle que, pour tout  $\alpha$ , nous ayons  $R_\alpha \subset R'_\alpha \subset R_\alpha$ . La proposition résulte alors des deux lemmes suivants.

**Lemme 11.31.** *Il existe  $c_1, c_2 > 0$  tels que, pour tout  $n \geq 0$  et tout évènement  $C$  non  $\mu$ -négligeable de la partition mesurable  $\bigvee_{i=0}^{n-1} \phi_M^{-1} \xi_0$ , nous ayons*

$$c_1 \lambda_+^{-n} \leq \mu(C) \leq c_2 \lambda_+^{-n} .$$

**Démonstration.** Notons  $m = m_s \otimes m_u$  la mesure de Lebesgue de  $\mathbb{R}^N = E^s \times E^u$ , qui s'identifie localement à  $\mu$  par la projection canonique  $\uparrow$ . Nous avons donc, pour tout rectangle  $R_\alpha = R_\alpha^s \times R_\alpha^u$ , l'égalité  $\mu(R_\alpha) = m_s(R_\alpha^s) m_u(R_\alpha^u)$ . Notons que puisque le rectangle  $R_\alpha$  est propre, les intérieurs de  $R_\alpha^s$  et  $R_\alpha^u$  sont non vides, donc  $m_s(R_\alpha^s) > 0$  et  $m_u(R_\alpha^u) > 0$ . De plus, pour tout  $w$  dans  $\Sigma_A$ , nous avons

$$\bigcap_{i=0}^{n-1} \phi^{-i}(R_{\omega_i}) = R_{\omega_0}^s \times \phi^{-(n-1)}(R_{\omega_{n-1}}^u) .$$

Donc, si  $C = \bigcap_{i=0}^{n-1} \phi^{-i}(R_{\omega_i})$ , alors  $\mu(C) = \lambda_+^{-n+1} m_s(R_{\omega_0}^s) m_u(R_{\omega_{n-1}}^u)$ . Il suffit par conséquent de prendre pour  $c_1$  le minimum des  $\lambda_+ m_s(R_\alpha^s) m_u(R_\beta^u)$  pour  $\alpha, \beta \in \mathcal{A}$  et pour  $c_2$  le maximum correspondant.  $\square$

**Lemme 11.32.** *Soient  $X$  un espace métrique compact,  $\phi$  un homéomorphisme de  $X$ ,  $\mu$  une mesure de probabilité borélienne  $\phi$ -invariante et ergodique sur  $X$ , et  $\xi_0$  une partition mesurable de  $X$  génératrice pour  $\phi$ . Supposons qu'il existe  $c_1, c_2, \lambda_+ > 0$  tels que, pour tout évènement  $C$  non  $\mu$ -négligeable de la partition  $\bigvee_{i=0}^{n-1} \phi^{-i} \xi_0$ , et pour tout  $n \geq 0$ , nous ayons*

$$c_1 \lambda_+^{-n} \leq \mu(C) \leq c_2 \lambda_+^{-n} .$$

Alors  $h_{\text{top}}(\phi) = h_\mu(\phi) = \log \lambda_+$ . En outre,  $\mu$  est l'unique mesure d'entropie maximale de  $\phi$  sur  $X$ .

**Démonstration.** Par la définition de l'entropie d'une partition mesurable  $\beta$  et puisque  $\sum_{B \in \beta} \mu(B) = 1$ , nous avons

$$H_\mu(\beta) \in [-\ln(\max\{\mu(B) : B \in \beta\}), -\ln(\min\{\mu(B) : B \in \beta\})] .$$

L'hypothèse assure donc que, pour tout  $n \geq 0$ , nous avons

$$H_\mu\left(\bigvee_{i=0}^{n-1} \phi^{-i} \xi_0\right) \in [-\ln(c_2 \lambda_+^{-n}), -\ln(c_1 \lambda_+^{-n})] .$$

Le théorème de Kolmogorov-Sinai donne alors

$$h_\mu(\phi) = h_\mu(\phi, \xi_0) = \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} \phi^{-i} \xi_0\right) = \ln(\lambda_+) .$$

D'après le principe variationnel (le théorème 11.11), il reste à montrer que si une mesure de probabilité borélienne  $\phi$ -invariante  $\mu'$  sur  $X$  vérifie  $h_{\mu'}(\phi) \geq h_\mu(\phi)$ , alors  $\mu' = \mu$ . Pour cela, il suffit de reprendre quasiment mot à mot la démonstration du théorème 11.24 b).  $\square$

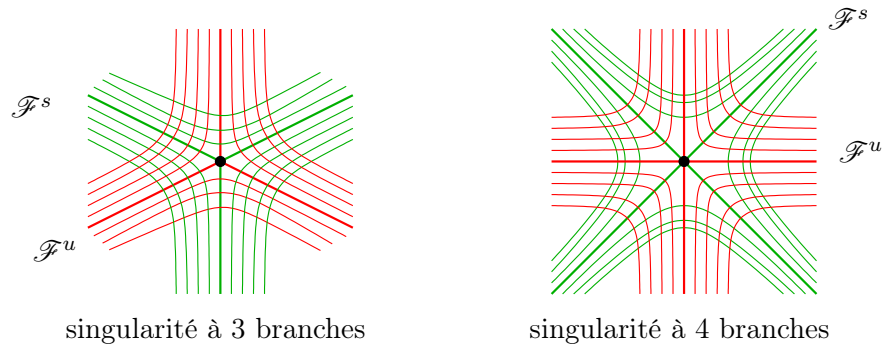
Reprenons la démonstration de la proposition 11.30. Nous avons vu que  $\phi_M$  est mélangeante, donc ergodique, pour la mesure de Haar  $\mu$  (voir le théorème 4.10). La partition  $\xi_0$  construite en début de démonstration est génératrice pour  $\phi_M$  (qui est inversible), car par

la démonstration de l’assertion **b)** du théorème 11.29, les évènements de  $\bigvee_{i=-n}^n \phi_M^{-i} \xi_0$  sont des rectangles d’intérieurs disjoints et de diamètre tendant vers 0, donc qui engendrent la  $\sigma$ -algèbre des boréliens. Le lemme 11.32 précédent permet alors de conclure la démonstration de la proposition 11.30.  $\square$

Ainsi, si  $N = 2$ , et si  $\lambda, \lambda'$  sont les deux valeurs propres d’une matrice hyperbolique  $M \in \text{GL}_2(\mathbb{Z})$  de taille 2, avec  $|\lambda| > 1 > |\lambda'|$ , nous avons

$$h_{\text{top}}(\phi_M) = \ln |\lambda| .$$

Ce résultat s’étend (voir [FLP]) aux homéomorphismes pseudo-Anosov, au sens ci-dessous.



Nous dirons qu’un homéomorphisme  $\phi$  préservant l’orientation d’une surface compacte connexe orientable (sans bord)  $S$  est *pseudo-Anosov* s’il existe  $\lambda > 1$  (appelé le *facteur de dilatation* de  $\phi$ ) et  $(\mathcal{F}^s, \mu^s)$  et  $(\mathcal{F}^u, \mu^u)$  deux feuilletages à singularités isolées de type selle à au moins trois branches, invariants par  $\phi$  et transverses, munis d’une mesure transverse de support total invariante par holonomie, telle que  $\phi_* \mu^s = \frac{1}{\lambda} \mu^s$  et  $\phi_* \mu^u = \lambda \mu^u$ . Nous ne donnons pas les définitions précises (voir loc. cit.), mais pour  $N = 2$  et  $\phi_M$  comme ci-dessus, les feuilletages (non singuliers) définis par les variétés stables et les variétés instables, avec pour mesure transverse pour un feuilletage la mesure de Lebesgue le long des arcs de feuilles de l’autre feuilletage, conviennent. Le calcul de l’entropie topologique d’un homéomorphisme pseudo-Anosov est analogue à celui des difféomorphismes Anosov linéaires  $\phi_M$  du tore  $\mathbb{T}^2$ . Elle est égale (voir loc. cit.) au logarithme de son facteur de dilatation :

$$h_{\text{top}}(\phi_M) = \ln \lambda .$$

Dans une autre direction, les résultats qui précèdent sont des cas particuliers du résultat suivant (qui étend des résultats de Sinai et Ruelle), pour lequel nous renvoyons à [Bow1, Bow3], avec en plus une extension aux mesures d’équilibre. L’ergodicité découle de l’exercice E.70, et la conjugaison (en tant que système dynamique mesuré) à un sous-décalage de type fini procède par un codage, en utilisant une construction de partition de Markov analogue à celle qui a été effectuée pour les automorphismes linéaires des tores à partir de la structure de produit local des pièces basiques (voir la partie 10.9).

**Théorème 11.33. (Théorème de Bowen)** *Soit  $\Lambda$  une pièce basique d’un  $C^1$ -difféomorphisme  $\phi$  Axiome A d’une variété lisse  $M$ . Alors  $\phi|_{\Lambda}$  admet une et une seule mesure d’entropie maximale  $\mu_{\Lambda}$ . Celle-ci est ergodique. Les points périodiques de  $\phi|_{\Lambda}$  s’équidistribuent*

vers  $\mu_\Lambda$  : si  $\text{Per}_N(\phi|_\Lambda) = \{x \in \Lambda : \phi^N(x) = x\}$  est l'ensemble des points périodiques de  $\phi|_\Lambda$  de période divisant  $N$ , et si  $\Delta_x$  est la masse de Dirac unité en  $x$ , alors

$$\frac{1}{\text{Card Per}_N(\phi|_\Lambda)} \sum_{x \in \text{Per}_N(\phi|_\Lambda)} \Delta_x \xrightarrow{*} \mu_\Lambda .$$

Le système dynamique mesuré  $(\Lambda, \mu_\Lambda, \phi|_\Lambda)$  est conjugué à une chaîne de Markov bilatère sur un alphabet fini, qui est mélangeante si  $\phi|_\Lambda$  est topologiquement mélangeante.  $\square$

## 11.5 Exercices

**Exercice E.68.** (1) Montrer que toute rotation du cercle est d'entropie topologique nulle. Plus généralement, soient  $(X, d)$  un espace métrique compact et  $\phi : X \rightarrow X$  une application 1-lispchitzienne (ce qui signifie que pour tous les  $x, y$  dans  $X$ , nous avons l'inégalité  $d(\phi(x), \phi(y)) \leq d(x, y)$ ). Montrer que  $h_{\text{top}}(\phi) = 0$ .

(2) Montrer que tout homéomorphisme  $\phi$  du cercle préservant l'orientation, de nombre de rotation irrationnel, est uniquement ergodique, est d'entropie topologique nulle, et admet une unique mesure d'entropie maximale.

**Exercice E.69.** Soient  $X$  et  $Y$  des espaces métriques compacts. Considérons des applications continues  $f : X \rightarrow X$ ,  $g : Y \rightarrow Y$  et  $\mathfrak{h} : X \rightarrow Y$  avec  $\mathfrak{h}$  surjective, telles que  $\mathfrak{h} \circ f = g \circ \mathfrak{h}$ .<sup>253</sup> Montrer que

$$h_{\text{top}}(g) \leq h_{\text{top}}(f) .$$

**Exercice E.70.** Soient  $X$  un espace métrique compact et  $\phi : X \rightarrow X$  une application continue. Montrer que si  $\phi$  admet une unique mesure d'entropie maximale  $\mu$ , alors  $\mu$  est un point extrémal du convexe  $\text{Prob}_\phi(X)$ , et donc que  $\phi$  est ergodique pour  $\mu$ .

**Exercice E.71.** Soient  $X$  le cercle  $\mathbb{R}/\mathbb{Z}$  et  $\phi_N : X \rightarrow X$  l'application définie par  $x \mapsto Nx$ , pour  $N$  dans  $\mathbb{N} - \{0, 1\}$ . Le but de cet exercice est, entre autre, de montrer que la mesure de Lebesgue de  $\mathbb{R}/\mathbb{Z}$  est l'unique mesure d'entropie maximale pour  $\phi_N$ .

- (1) Soit  $\mu$  une mesure de probabilité borélienne invariante par  $\phi_N$  sur le cercle  $X$ . Montrer que  $h_\mu(\phi_N) \leq \ln N$ .<sup>254</sup>
- (2) Montrer qu'il y a égalité si et seulement si  $\mu$  est la mesure de Lebesgue  $\lambda$  sur  $X$ .
- (3) Calculer l'entropie topologique de  $\phi_N$ .

**Exercice E.72.** Soient  $\mathcal{A}$  un alphabet fini de cardinal  $N$ , et  $(\Sigma_+ = \mathcal{A}^\mathbb{N}, \sigma_+)$  le système de Bernoulli unilatère sur  $\mathcal{A}$ . Notons  $\nu$  l'équiprobabilité sur  $\mathcal{A}$ . Le but de cet exercice est, entre autre, de montrer que la mesure produit  $\nu^\mathbb{N}$  est l'unique mesure d'entropie maximale pour le décalage de Bernoulli  $\sigma_+$ .

- (1) Soit  $\mu$  une mesure de probabilité borélienne invariante par  $\sigma_+$  sur  $\Sigma_+$ . Montrer que  $h_\mu(\sigma_+) \leq \ln N$ .
- (2) Montrer qu'il y a égalité si et seulement si  $\mu$  est la mesure produit  $\nu^\mathbb{N}$  sur  $X$ .

253. L'application  $\mathfrak{h}$  est donc une semi-conjugaison entre les systèmes dynamiques topologiques  $(X, f)$  et  $(Y, g)$ .

254. Plus généralement, montrer que si  $(X, \mathcal{B}, \mu, \phi)$  est un système dynamique mesuré, où  $X$  est un espace métrique compact et  $\mu$  une mesure de probabilité, et si  $X$  admet une partition mesurable génératrice pour  $\phi$  de cardinal  $N$ , alors  $h_\mu(\phi) \leq \ln N$ .

- (3) Calculer l'entropie topologique de  $\sigma_+$ .
- (4) Montrer que la mesure produit  $\nu^{\mathbb{Z}}$  est l'unique mesure d'entropie maximale pour le système de Bernoulli bilatère  $(\Sigma = \mathcal{A}^{\mathbb{Z}}, \sigma)$ .

**Exercice E.73.** Soient  $(\Sigma_A, \sigma_A)$  un sous-décalage de type fini bilatère de matrice d'incidence  $A$  apériodique et  $P_n = \{\omega \in \Sigma_A : \sigma_A^n(\omega) = \omega\}$  l'ensemble des points périodiques de  $\sigma_A$  de période divisant  $n$ . Montrer que, quand  $n$  tend vers  $+\infty$ , nous avons

$$\text{Card } P_n \sim e^{n h_{\text{top}}(\sigma_A)} .$$

**Exercice E.74.** Soient  $M \in \text{SL}_2(\mathbb{Z})$  une matrice hyperbolique, de valeurs propres  $\lambda_1, \lambda_2$ , et  $\phi_M : \mathbb{T}^2 \rightarrow \mathbb{T}^2$  l'automorphisme linéaire hyperbolique du tore  $\mathbb{T}^2$  associé. Pour tout  $n \in \mathbb{N} - \{0\}$ , notons  $P_n = \{x \in \mathbb{T}^2 : \phi_M^n(x) = x\}$  l'ensemble des points périodiques de  $\phi_M$  de période divisant  $n$ .

- (1) Montrer qu'un point de  $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$  est périodique pour  $\phi_M$  si et seulement s'il est rationnel (c'est-à-dire appartient à  $\mathbb{Q}^2/\mathbb{Z}^2$ ).
- (2) Montrer que  $\text{Card } P_n = |\lambda_1^n + \lambda_2^n - 2|$ .
- (3) Montrer que, quand  $n$  tend vers  $+\infty$ , nous avons

$$\text{Card } P_n \sim e^{n h_{\text{top}}(\phi_M)} .$$

# Index

- $\sigma$ -additive, 20
- algèbre
  - de Boole, 13
  - séparante, 85
- alphabet, 9, 124, 284
- angle, 160
- application
  - à variation bornée, 163
  - analytique réelle, 169
  - caractéristique, 30
  - contractante, 171
  - de doublement de l'angle, 148, 159
  - de Gauss, 131
  - différentiable, 84
  - dilatante, 92
  - du chat d'Arnold, 217
  - du fer à cheval de Smale, 231
  - équicontinues, 152
  - expansive, 242
    - positivement, 242
  - höldérienne, 70
  - invariante, 12
  - lipschitzienne, 212
  - mélangeante, 66
  - monotone, 164
  - presque invariante, 12
  - propre, 22
  - préservant la mesure, 6
  - résolvante, 262
  - s'annulant à l'infini, 21
  - semi-continue supérieurement, 149
  - sous-harmonique, 59
  - strictement dilatante, 170
  - trigonométrique, 84
  - uniformément continue, 211
  - étagée, 67
- atome, 98
- attracteur de Smale, 234
- attractif, 139
- base dénombrable d'ouverts, 56
- bassin
  - d'attraction, 234
  - de répulsion, 234
- bloc de Jordan, 210
- borné, 222
- champ
  - de cônes, 235
  - de vecteurs, 224
- chat d'Arnold, 217
- cobord, 55
- cocycle
  - linéaire, 42
  - multiplicatif, 42
- codage, 148, 284
- coefficient
  - de corrélation, 69
  - de Fourier, 72, 84, 85
  - matriciel, 179
- compactifié d'Alexandrov, 139
- complet, 17
- complexifié, 263
- condition de Renyi, 133
- condition diophantienne, 170
- conjugaison, 8
  - ensembliste, 7
- conjugués, 8
  - ensemblissement, 7
- constante
  - d'expansivité, 242
  - d'hyperbolicité, 209
  - de Lipschitz, 212
- convergence faible, 183
- courbe modulaire, 197
- critère
  - de Weyl, 88
- critère de Weyl, 288
- crochet de Bowen, 257
- cylindre, 10, 15, 126
- décalage, 15
  - de Bernoulli, 10
  - de Markov, 126
  - sous-décalage de type fini, 285
- décomposition
  - de Cartan, 45, 183
  - de Hopf, 27
  - en valeurs singulières, 45
- décroissance des corrélations, 69
  - exponentielle, 69
  - polynomiale, 69
  - rapide, 69
- degré, 159
- demi-plan supérieur, 186
- $\epsilon$ -dense, 275
- dérivée de Radon-Nykodym, 55
- dérivées partielles, 84
- développement en fraction continue, 130
- différence symétrique, 12
- différentielle, 219
- difféomorphisme
  - Anosov, 229
  - Axiome A, 229



- de Kupka-Smale, 262
- dilatante, 170
- diophantien, 170
- distance
  - de Hilbert, 123
  - de pistage, 277
  - de Rokhlin, 112
  - hyperbolique, 191
  - invariante à gauche, 174
- domaine fondamental, 190
  - strict, 191
- dual topologique, 21
- dynamique Nord-Sud, 140
- endomorphisme linéaire hyperbolique, 208
- ensemble
  - $\alpha$ -limite, 142
  - non errant, 143
  - $\omega$ -limite, 142
- entropie, 102, 107, 276
  - mesurée, 98, 107
  - métrique, 98, 107
  - relative, 102
  - relativement à une partition, 107
  - relativement à un recouvrement, 276
  - topologique, 276, 278
- équidistribution, 287
- équirépartition, 88
- ergodique, 29
  - uniquement, 86
- errant, 24, 143
- escalier de Lebesgue, 168
- espace
  - à l'infini, 192
  - de Baire, 144
  - de Cantor, 10
    - triadique, 10
  - de drapeaux, 46
  - des drapeaux, 46
  - des phases, 6
  - vectériel topologique, 20
- espérance, 17
  - conditionnelle, 32
- évènement, 98
- expansif, 242
  - positivement, 242
- exposant
  - de Lyapounov, 47
    - maximal, 44
    - minimal, 47
- extension naturelle, 10
- facteur, 8
  - de Pinsker, 118
- fer à cheval de Smale, 231
- fibre, 221
- fibré vectoriel
  - morphisme, 221
- fibré tangent, 186
  - unitaire, 187
- fibré vectoriel, 221
  - normé, 222
  - rang, 221
  - restriction, 222
  - somme directe, 222
- fixateur, 179
- flot, 7
  - d'Anosov, 230
  - géodésique, 195, 196
  - horocyclique, 195, 196
    - instable, 197
    - stable, 196
- fonction
  - de hauteur, 57
  - de Lyapounov, 47
  - indicatrice, 30
- fraction continue, 130
- frontière, 146
- géodésique hyperbolique, 192
- groupe
  - modulaire, 189
  - topologique, 15
  - unimodulaire, 175
  - unitaire, 178
  - à un paramètre, 6
- $H_\mu(\alpha)$ , 102
- $H_\mu(\alpha|\beta)$ , 102
- $h(\phi, \mathcal{U})$ , 276
- $h_\mu(\phi)$ , 107
- $h_\mu(\phi, \alpha)$ , 107
- $h_{\text{top}}(\phi)$ , 276
- hamiltonien, 19
- homéomorphisme pseudo-Anosov, 301
- horicycle, 194
- horoboule, 204
- horocycle, 194
  - instable, 194
  - stable, 194
- hyperbolique, 208, 215, 229
- indépendance, 100
- information, 98, 104
  - relative, 104
- instable, 208, 225, 230
- intégrale première, 19

- invariante, 11, 208
  - globalement, 11
  - positivement, 11
  - presque, 12
- inégalité de Hölder, 80
- joint, 99, 274
- K-système, 118
- lemme
  - de Fekete, 36
  - de fermeture d'Anosov, 245
  - de Lebesgue, 277
  - de pistage, 243
    - linéaire, 260
  - ergodique maximal, 32
- Liouville, 170
- localement maximal, 258
- loi, 17
  - forte des grands nombres, 34
- longueur
  - d'arc, 191
  - hyperbolique, 191
- matrice
  - apériodique, 122
  - d'incidence, 285, 298
    - associée, 124
  - de transition, 124
  - irréductible, 122
    - période d'une, 135
  - stochastique, 124
- mélange
  - d'ordre  $k$ , 76
  - faible, 67, 76
  - fort, 66
  - mesuré, 66, 148
    - effectif, 69
    - exponentiel, 69, 148
    - polynomial, 69
    - rapide, 69
  - multiple, 76
  - topologique, 145
- mesure, 56
  - absolument continue, 55
  - complète, 6
  - d'entropie maximale, 283
  - de Liouville, 189
  - de Gauss, 132
  - de Haar, 15, 84, 175, 177, 186
    - normalisée, 15
  - de Lebesgue, 199
  - de Liouville, 18, 187
- de Markov
  - bilatère, 126
  - unilatère, 126
- de Parry, 290
- ergodique, 29
- image, 6
- invariante, 6
- mélangeante, 66
- $\sigma$ -finie, 13
- signée, 20
- spectrale, 74
- sphérique, 187
- stationnaire, 125
- étrangères, 55

- minimal, 146
  - exceptionnel, 165
- moins fin, 275
- moins fine, 99
- mot, 10
- moyenne
  - de Birkhoff, 31
  - orbitale, 31
  - spatiale, 31
  - temporelle, 31
- $N'_d(\epsilon)$ , 275
- $N_d(\epsilon)$ , 275
- neutre, 230
- nombre
  - de translation, 160
  - de Liouville, 170
  - de rotation, 160
  - de type Roth, 170
  - diophantien, 170
- non errant, 143
- norme, 222, 224
  - adaptée, 209, 227
  - d'opérateur, 208, 222
  - de Sasaki, 230
  - de Sobolev, 70
  - duale, 21
  - hyperbolique, 187
  - höldérienne, 70
  - matricielle, 44
  - riemannienne, 239
  - uniforme, 21
  - équivalentes, 222, 224
- opérateur
  - de Koopman, 70
  - de transfert, 95, 148
- orbite, 8
  - complète, 8

- positive, 8
- périodique, 8
- ordre de Sharkovsky, 140
- paramétrage local, 6
- partie
  - $\epsilon$ -dense, 275
  - entière, 88
  - fractionnaire, 88, 129, 288
  - hyperbolique, 225
  - localement maximale, 258
  - minimale, 146
  - $\epsilon$ -séparée, 275
- partition
  - de Markov, 297
  - image réciproque, 101
  - indépendantes, 100
  - jointe, 99
  - mesurable, 98
    - génératrice, 107
    - moins fine, 99
  - presque partout moins fine, 101
  - presque partout égales, 100
  - triviale, 98
- période, 8, 9
- périodique, 8
  - attractif, 139
  - répulsif, 139
- phénomène de Mautner, 180
- $\pi$ -système, 13
- pistage, 243, 246
  - linéaire, 260
- pièce basique, 259
- plan hyperbolique réel, 187
- point
  - à l'infini, 192
  - errant, 143
  - extremal, 20
  - fixe hyperbolique, 215
  - homocline, 262
  - récurrent
    - négativement, 142
    - positivement, 142
- positivement minimal, 146
- principe variationnel, 279
- produit
  - aléatoire de matrices, 43
  - dynamique de matrices, 42
- produit local, 257, 295
- projection
  - canonique, 84
  - de Cartan, 45
  - stéréographique, 140
- propre, 22
- préserver la mesure, 6
- pseudo-orbite, 243, 246, 260
- raffinement, 99
- rayon géodésique, 192
- rayon spectral, 262, 286
- recouvrement
  - générateur, 278
  - joint, 274
  - moins fin, 275
- rectangle, 296
  - propre, 297
- récurrent
  - négativement, 142
  - positivement, 24, 142
- relèvement, 218
- relevé, 159
- représentation
  - quasi-régulière, 178
  - unitaire, 178
- répulsif, 139
- réseau, 177, 186
  - cocompact, 177, 186
  - uniforme, 177, 186
- rotation, 9
- section, 222
  - ensembliste, 222
- semi-algèbre de Boole, 14
- semi-anneau de Boole, 13
- semi-conjugaison, 7
  - ensembliste, 7
- semi-conjugués
  - ensemblissement, 7
- séparable, 56
- séparante, 85
- $\epsilon$ -séparé, 275
- sous-additive, 35
- sous-décalage de type fini
  - bilatère, 285
  - unilatère, 285
- sous-espace
  - instable, 208
  - stable, 208
- sous-fibré, 221
  - instable, 225, 230
  - neutre, 230
  - stable, 225, 230
- sous-rectangle
  - instable, 297
  - stable, 297
- spectre, 210, 262, 264

- de Lyapounov, 47
- ponctuel, 71
- stable, 208, 225, 230
- structurellement stable, 261
  - linéairement, 260
- suite
  - équirépartie, 88
- suite de Farey, 289
- support, 56
- système
  - de Bernoulli
    - bilatère, 9
    - unilatère, 9
  - de Kronecker, 15
  - de Markov, 285
    - bilatère, 126
    - unilatère, 126
  - hamiltonien, 19
- système dynamique, 6
  - à temps continu, 6
  - à temps discret, 6
  - conjugué, 8
  - de Kolmogorov, 118
  - différentiable, 7
    - de classe  $C^k$ , 7
  - ergodique, 29
  - faiblement mélangeant, 76
  - induit, 56
  - inversible, 7
  - mesurable, 6
  - mesuré, 6
    - complètement dissipatif, 25
    - conservatif, 25
  - mélangeant, 66
  - Nord-Sud, 140
  - probabilisé, 6
  - produit croisé, 57
  - structurellement stable, 261
  - topologique, 6
    - expansif, 242
    - non errant, 143
  - totalement ergodique, 58
- temps de retour
  - $k$ -ème, 26
  - premier, 26
- théorème
  - d'Abramov, 118
  - d'Anosov, 220
  - d'Arzela-Ascoli, 152
  - d'Ensaï, 58
  - d'extension de Carathéodory, 13
  - d'inversion locale, 271
  - d'isomorphisme d'Ornstein, 111
  - d'Oseledets, 46
  - de Bochner, 73
  - de Baire, 144
  - de Banach, 262
  - de Banach-Alaoglu, 22
  - de Birkhoff, 31
  - de Bowen, 301
  - de Dani-Furstenberg, 198
  - de Denjoy, 163
  - de Dirichlet, 130
  - de dualité de Riesz, 183
  - de décomposition de Jordan, 211
  - de Furstenberg-Kesten, 44
  - de Furstenberg-Weiss, 149
  - de Gauss, 132
  - de Grobman-Hartman, 215
  - de Howe-Moore, 182
  - de Kac, 26
  - de Kingman, 35
  - de Kolmogorov, 34
  - de Kolmogorov-Sinaï, 108
  - de Krein-Milman, 22
  - de Krylov-Bogolyubov, 20
  - de Liouville, 18
  - de Perron-Frobenius, 122
  - de Radon-Nykodym, 55
  - de représentation de Riesz, 21
  - de récurrence
    - de Birkhoff, 147
    - de Poincaré, 25
  - de Sard, 261
  - de semi-conjugaison de Poincaré, 164
  - de Sharkovsky, 140
  - de Stone-Weierstrass, 85
  - de van der Waerden, 151
  - de Weyl, 88, 90
  - du facteur de Sinai, 111
  - du point fixe de Banach, 173
  - du principe variationnel, 279
  - ergodique
    - d'Oseledets, 46
    - de Birkhoff, 31
    - de Kingman, 35
- topologie
  - $C^1$ , 219, 238
  - $C^k$ , 238
  - faible-étoile, 21
  - vague, 21
- tore, 84
- totalement discontinu, 10
- totalement ergodique, 58
- tour de Kakutani, 57

- transformation, 7
  - de Fourier, 72
  - de Gauss, 131
  - de Kakutani, 57
  - ergodique, 29
  - induite, 56
  - totale<sup>ment</sup> ergodique, 58
- transient, 24
- transitif, 144
  - positivement, 144
  - topologiquement, 144
- translation, 84
  - à droite, 15
  - à gauche, 15
- tribu des invariants, 32
- trivialisation locale, 221
  
- uniformément continue, 211
- unimodulaire, 175
  
- valeur
  - critique, 261
  - régulière, 261
- variable aléatoire, 17
  - indépendantes, 17
  - uniformément distribuées, 17
- variation, 20
  - totale, 21
- variation bornée, 163
- variété
  - instable, 255, 295
    - locale, 251, 294
  - stable, 255, 295
    - locale, 251, 294
- volume hyperbolique, 187

## Références

- [Aar] J. Aaronson. *An introduction to infinite ergodic theory*. Math. Surv. Mono. **50**, Amer. Math. Soc. 1997.
- [AkS] M. Aka and U. Shapira. *On the evolution of continued fractions in a fixed quadratic field*. J. Anal. Math. **134** (2018) 335–397.
- [AIB] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis, A hitchhiker’s guide*. 2nd ed., Springer Verlag, 1999.
- [Arn] L. Arnold. *Random Dynamical Systems*. Springer Mono. Math., Springer Verlag, 2003.
- [AP] D. K. Arrowsmith and C. M. Place. *An introduction to Dynamical systems*. Cambridge Univ. Press, 1990.
- [AvB] A. Avila and J. Bochi. *On the subadditive ergodic theorem*. <https://personal.science.psu.edu/jzd5895/docs/kingbirk.pdf>.
- [Bal] V. Baladi. *Dynamical zeta functions and dynamical determinants for hyperbolic maps. A functional approach*. Erg. Math. Grenz. **68**. Springer Verlag, 2018.
- [BaL] V. Baladi and C. Liverani. *Exponential decay of correlations for piecewise cone hyperbolic contact flows*. Comm. Math. Phys. **314** (2012) 689–773.
- [BeHV] B. Bekka, P. de la Harpe, and A. Valette. *Kazhdan’s property T*. New Math. Mono. **11**, Cambridge Univ. Press, 2008.
- [BekM] B. Bekka and M. Mayer. *Ergodic theory and topological dynamics of group actions on homogeneous spaces*. London Math. Soc. Lect. Note Ser. **269**, Cambridge Univ. Press, 2000.
- [BenP] Y. Benoist and F. Paulin. *Systèmes dynamiques élémentaires*. Notes de cours 2003, Ecole Normale Supérieure, <http://www.math.u-psud.fr/~paulin/notescours>.
- [Bil] P. Billingsley. *Convergence of probability measures*. Wiley, 1968.
- [Bor] A. Borel. *Introduction aux groupes arithmétiques*. Hermann, 1969.
- [BorHC] A. Borel and Harish-Chandra. *Arithmetic subgroups of algebraic groups*. Ann. of Math. **75** (1962) 485–535.
- [Bou] N. Bourbaki. *Topologie générale*. chap. 1 à 4, Hermann, 1971.
- [Bow1] R. Bowen. *Markov Partitions for Axiom A Diffeomorphisms*. Amer. J. Math. **92** (1970) 725–747.
- [Bow2] R. Bowen. *Maximizing entropy for a hyperbolic flow*. Math. Syst. Theo. **7** (1974) 300–303.
- [Bow3] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Lect. Notes Math. **470**, Springer Verlag, 1975.
- [Bre] H. Brezis. *Analyse fonctionnelle*. Masson, 1983.
- [BriD] J.-Y. Briand and J. Duval. *Deux caractérisations de la mesure d’équilibre d’un endomorphisme de  $\mathbb{P}_k(\mathbb{C})$* . Publ. Math. Inst. Hautes Études Sci. **93** (2001) 145–159.
- [BrPP] A. Broise-Alamichel, J. Parkkonen, and F. Paulin. *Equidistribution and counting under equilibrium states in negative curvature and trees. Applications to non-Archimedean Diophantine approximation*. With an Appendix by J. Buzzi. Prog. Math. **329**, Birkhäuser, 2019.
- [Bur] D. Burguet. *Existence of measures of maximal entropy for  $C^r$  interval maps*. Proc. Amer. Math. Soc. **142** (2014) 957–968.
- [BurH] K. Burns and B. Hasselblatt. *The Sharkovsky theorem : a natural direct proof*. Amer. Math. Monthly **118** (2011) 229–244.

- [BurMMW] K. Burns, H. Masur, C. Matheus, and A. Wilkinson. *Rates of mixing for the Weil-Petersson geodesic flow : exponential mixing in exceptional moduli spaces*. *Geom. Funct. Anal.* **27** (2017) 240–288.
- [Buz1] J. Buzzi. *Maximal entropy measures for piecewise affine surface homeomorphisms*. *Erg. Theo. Dyn. Syst.* **29** (2009) 1723–1763.
- [Buz2] J. Buzzi.  *$C^r$  surface diffeomorphisms with no maximal entropy measure*. *Erg. Theo. Dyn. Syst.* **34** (2014) 1770–1793, 2014.
- [BuzCS] J. Buzzi, S. Crovisier, and O. Sarig. *Measures of maximal entropy for surface diffeomorphisms*. Prépublication [arXiv:1811.02240](https://arxiv.org/abs/1811.02240).
- [BuzS] J. Buzzi and O. Sarig. *Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps*. *Erg. Theo. Dyn. Syst.* **23** (2003) 1383–1400.
- [Cal] D. Calegari. *Stable commutator length*. *MSJ memoirs* **20**, Math. Soc. Japan, 2009.
- [CanC1] A. Candel and L. Conlon. *Foliations I*. *Grad. Stud. Math.* **23**, Amer. Math. Soc. 2000.
- [CanC2] A. Candel and L. Conlon. *Foliations II*. *Grad. Stud. Math.* **60**, Amer. Math. Soc. 2003.
- [Car] H. Cartan. *Cours de calcul différentiel*. Hermann, 2nde éd. 1977.
- [Coh] D. Cohn. *Measure theory*. Birkhäuser, 1980.
- [CorFS] I. P. Cornfeld and S. V. Fomin and Ya. G. Sinai. *Ergodic theory*. *Grund. Math. Wiss.* **245**, Springer-Verlag, 1982.
- [Cou] Y. Coudène. *Théorie ergodique et systèmes dynamiques*. *Savoirs actuels*, CNRS Editions-EDP Sciences, 2013.
- [Dal] F. Dal’Bo. *Trajectoires géodésiques et horocycliques*. EDP Sciences, 2007.
- [Die1] J. Dieudonné. *Éléments d’analyse t. 1 : fondements de l’analyse moderne*. Gauthier-Villars, 1971.
- [Die2] J. Dieudonné. *Éléments d’analyse t. 2 : chap. XII à XV*. Gauthier-Villars, 1974.
- [DiS] T.-C. Dinh and N. Sibony. *Equidistribution problems in complex dynamics of higher dimension*. *Internat. J. Math.* **28** (2017) 1750057.
- [DoS] T. Downarowicz and J. Serafin. *A short proof of Orstein theorem*. *Erg. Theo. Dyn. Syst.* **32** (2012) 587-597.
- [Dug] J. Dugundji. *Topology*. Wm. C. Brown, 1989.
- [EL] M. Einsiedler and E. Lindenstrauss. *Diagonal actions on locally homogeneous spaces*. In "Homogeneous flows, moduli spaces and arithmetic", pp 155–241, *Clay Math. Proc.* **10**, Amer. Math. Soc. 2010.
- [EW] M. Einsiedler and T. Ward. *Ergodic theory with a view towards number theory*. *Grad. Texts Math.* **259**, Springer Verlag, 2011.
- [FLP] A. Fathi, F. Laudenbach, and V. Poenaru. *Travaux de Thurston sur les surfaces*. *Astérisque* **66-67**, Soc. Math. France 1979.
- [FuK] H. Furstenberg, H. Kesten. *Products of Random Matrices*. *Ann. Math. Stat.* **31** (1960) 457–469.
- [GaHL] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian geometry*. Springer Verlag, 1990.
- [Gei] H. Geiges. *A introduction to contact topology*. *Cambridge Stud. Adv. Math.* **109**, Cambridge Univ. Press, 2008.
- [Ghy] E. Ghys. *Groups acting on the circle*. *Enseign. Math.* **47** (2001) 329–407.
- [God1] C. Godbillon. *Éléments de topologie algébrique*. Hermann, 1971.
- [God2] C. Godbillon. *Feuilletages. Études géométriques*. *Prog. Math* **98**, Birkhäuser, 1991.

- [Gro] M. Gromov. *Groups of polynomial growth and expanding maps*. Pub. Math. I.H.É.S. **53** (1981) 53–78.
- [Gur] B. Gurevich. *Topological entropy for denumerable Markov chains*. Soviet Math. Dokl. **10** (1969) 911–915.
- [Har] P. de la Harpe, *On Hilbert’s metric for simplices*, pp 97–119, Vol. 1 de “Geometric group theory”, A. Niblo, M. Roller eds, LMS LNS **182**, Cambridge Univ. Press, 1993.
- [Hat] A. Hatcher. *Algebraic topology*. Cambridge Univ. Press, 2002.  
<http://pi.math.cornell.edu/~hatcher/AT/ATpage.html>.
- [Her] M. Herman. *Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations*. Publ. Math. I.H.É.S. **49** (1979) 5–233.
- [Hir] M. Hirsch. *Differential topology*. Grad. Texts Math. **33**, Springer Verlag, 1976.
- [HT] R. Howe and E. C. Tan. *Non-abelian harmonic analysis : application of  $SL(2, \mathbb{R})$* . Universitext, Springer Verlag, 1992.
- [IK] H. Iwaniec and E. Kowalski. *Analytic number theory*. Colloquium Pub. **53**, Amer. Math. Soc. 2004.
- [KM] A. Karlsson and G. Margulis. *A Multiplicative Ergodic Theorem and Nonpositively Curved Spaces*. Commun. Math. Phys. **208** (1999) 107–123.
- [Kat] S. Katok. *Fuchsian groups*. Univ. Chicago Press, 1992.
- [KH] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*. Ency. Math. App. **54**, Camb. Univ. Press, 1995.
- [Khi] A. Y. Khinchin. *Continued fractions*. Univ. Chicago Press, 1964.
- [Kit] B. P. Kitchens. *Symbolic Dynamics : One-sided, Two-sided and Countable State Markov Shifts*. Universitext, Springer Verlag, 1998.
- [Kri] J.L. Krivine. *Théorie axiomatique des ensembles*. PUF, 1969.
- [Laf] J. Lafontaine. *Introduction aux variétés différentielles*. Press. Univ. Grenoble, 1996.
- [LeG] J.-F. Le Gall. *Intégration, probabilités et processus aléatoires*. Note de cours, Ecole Normale Supérieure, <http://www.math.u-psud.fr/~jfllegall/IPPA2.pdf>.
- [LB] P. Lévy-Bruhl. *Introduction à la théorie spectrale*. Dunod, 2003.
- [LY] T.-Y. Li and J. A. Yorke. *Period three implies chaos*. Amer. Math. Monthly **82** (1975) 985–992.
- [Man] R. Mané. *On the uniqueness of the maximizing measure for rational maps*. Bol. Soc. Brasil. Mat. **14** (1983) 27–43.
- [Mar1] G. Margulis. *Oppenheim conjecture*. In “Fields medalists lectures” 272–327, World Scien. Ser. 20th Cent. Math. **5**, World Scientific 1997.
- [Mar2] G. Margulis. *Diophantine approximation, lattices and flows on homogeneous spaces*. In “A panorama of number theory or the view from Baker’s garden” (Zürich, 1999), 280–310, Cambridge Univ. Press, 2002.
- [Mar3] G. Margulis. *Random walks on the space of lattices and the finiteness of covolumes of arithmetic subgroups*. In "Algebraic groups and arithmetic", 409–425, Tata Inst. Fund. Res., 2004.
- [Mel] I. Melbourne. *Superpolynomial and polynomial mixing for semiflows and flows*. Nonlinearity **31** (2018) 268–316.
- [Mil] J. Milnor. *Topology from the differentiable viewpoint*. Univ. Press Virginia, Charlottesville, 1965.
- [MT] N. Mneimné and F. Testard. *Introduction à la théorie des groupes de Lie classiques*. Hermann, 1986.



- [Nau] F. Naud. *On the rate of mixing of circle extensions of Anosov maps*. Spectral Theo. **9** (2019) 791–824.
- [Neve] J. Neveu. *Calcul des probabilités*. Masson, 1980.
- [Nevi] E. Neville. *The structure of Farey series*. Proc. London Math. Soc. **51** (1949) 132–144.
- [New] S. Newhouse. *Continuity properties of entropy*. Ann. of Math. **129** (1989) 215–235.
- [PaP1] J. Parkkonen and F. Paulin. *On the arithmetic of cross-ratios and generalised Mertens' formulas*. Numéro Spécial "Aux croisements de la géométrie hyperbolique et de l'arithmétique", F. Dal'Bo, C. Lecuire eds, Ann. Fac. Scien. Toulouse **23** (2014) 967–1022.
- [PaP2] J. Parkkonen and F. Paulin. *Dentelles et flocons de neige arithmétiques*. Images des Mathématiques, CNRS, 2015.
- [Pau1] F. Paulin. *De la géométrie et de la dynamique de  $SL_n(\mathbb{R})$  et  $SO_n(\mathbb{R})$* . Dans "Sur la dynamique des groupes de matrices et applications arithmétiques" (avec Gilles Courtois et Françoise Dal'Bo), Journées X-UPS 2007, N. Berline, A. Plagne, C. Sabbagh eds., Editions Ecole Polytechnique, 2007.
- [Pau2] F. Paulin. *Géométrie différentielle élémentaire*. Notes de cours de première année de master, Ecole Normale Supérieure, 2007, voir [https://www.imo.universite-paris-saclay.fr/~paulin/notescours/liste\\_notescours.html](https://www.imo.universite-paris-saclay.fr/~paulin/notescours/liste_notescours.html).
- [Pau3] F. Paulin. *Topologie, analyse, calcul différentiel*. Notes de cours de troisième année de licence, Ecole Normale Supérieure, 2009, voir [http://www.math.u-psud.fr/~paulin/notescours/cours\\_analyseI.pdf](http://www.math.u-psud.fr/~paulin/notescours/cours_analyseI.pdf).
- [Pau4] F. Paulin. *Topologie algébrique élémentaire*. Notes de cours de première année de master, Ecole Normale Supérieure, 2010, voir [http://www.math.u-psud.fr/~paulin/notescours/cours\\_topoalg.pdf](http://www.math.u-psud.fr/~paulin/notescours/cours_topoalg.pdf).
- [Pau5] F. Paulin. *Groupes et géométries*. Notes de cours de seconde année de master, Université Paris-Sud, 2014, voir [http://www.math.u-psud.fr/~paulin/notescours/cours\\_georiem.pdf](http://www.math.u-psud.fr/~paulin/notescours/cours_georiem.pdf).
- [Pau6] F. Paulin. *Théorie spectrale et analyse harmonique*. Notes de cours de seconde année de magistère, Université Paris-Sud, 2019, voir [http://www.math.u-psud.fr/~paulin/notescours/cours\\_magistere2.pdf](http://www.math.u-psud.fr/~paulin/notescours/cours_magistere2.pdf).
- [Pau7] F. Paulin. *Algèbre et géométrie euclidienne et hermitienne*. Notes de cours de seconde année de master, Université Paris-Saclay, 2020, voir [https://www.imo.universite-paris-saclay.fr/~paulin/notescours/cours\\_Agregation.pdf](https://www.imo.universite-paris-saclay.fr/~paulin/notescours/cours_Agregation.pdf).
- [Pau8] F. Paulin. *Introduction topologique à la géométrie*. Notes de cours de première année de master, Université Paris-Saclay, 2021, voir [https://www.imo.universite-paris-saclay.fr/~paulin/notescours/cours\\_GeometrieM1Orsay.pdf](https://www.imo.universite-paris-saclay.fr/~paulin/notescours/cours_GeometrieM1Orsay.pdf).
- [Pet] K. Petersen. *Ergodic theory*. Cambridge Stud. Adv. Math. **2**, Corr. ed, Cambridge Univ. Press, 1989.
- [Pol] M. Pollicott. *On the rate of mixing of Axiom A flows*. Invent. Math. **81** (1985) 413–426.
- [PY] M. Pollicott and M. Yuri. *Dynamical systems and ergodic theory*. London Math. Soc. Stud. Texts **40**, Cambridge Math. Soc. 1998.
- [Rat] M. Ratner. *The rate of mixing for geodesic and horocycle flows*. Erg. Theo. Dyn. Syst. **7** (1987) 267–288.
- [Rud1] W. Rudin. *Fourier analysis on groups*. Wiley 1990.
- [Rud2] W. Rudin. *Functional Analysis*. 2nd ed., McGraw-Hill, 1991.
- [Sar1] O. Sarig. *Thermodynamic formalism for countable Markov shifts*. Erg. Theo. Dyn. Syst. **19** (1999) 1565–1593.
- [Sar3] O. Sarig. *Subexponential decay of correlations*. Invent. Math. **150** (2002) 629–653.

- [Sar4] O. Sarig. *Existence of Gibbs measures for countable Markov shifts*. Proc. Amer. Math. Soc. **131** (2003) 1751–1758.
- [Sar5] O. Sarig. *Thermodynamic formalism for countable Markov shifts*. Proc. Symp. Pure Math. **89** (2015) 81–117, Amer. Math. Soc.
- [Ser] J.-P. Serre. *Cours d'arithmétique*. Press. Univ. France, Paris, 1970.
- [Shu] M. Shub. *Stabilité globale des systèmes dynamiques*. Astérisque **56**, Soc. Math. France, 1978.
- [Sma] S. Smale. *Differentiable dynamical systems*. Bull. Amer. Math. Soc. **73** (1967) 747–817.
- [Spa] E. Spanier. *Algebraic topology*. Tata McGraw-Hill, 1981.
- [Ste] J. M. Steele. *Kingman's subadditive ergodic theorem*. Ann. Inst. Poincaré **25** (1989) 93–98.
- [Tak] H. Takahasi. *Uniqueness of minimizer for countable Markov shifts and equidistribution of periodic points*. J. Stat. Phys. **181** (2020) 2415–2431.
- [Tsu] M. Tsujii. *Exponential mixing for generic volume-preserving Anosov flows in dimension three*. J. Math. Soc. Japan **70** (2018) 757–821.
- [Via] M. Viana. *Lectures on Lyapunov Exponents*. Cambridge Univ. Press, 2014.
- [Wal] P. Walters. *An introduction to ergodic theory*. Grad. Texts Math. **79**, Springer Verlag, 1982.
- [Wei] A. Weil. *L'intégration dans les groupes topologiques et ses applications*. Hermann, 1965.
- [Woe] W. Woess. *Denumerable Markov chains*. EMS Textbooks in Math, Europ. Math. Soc. 2009.
- [Yoc1] J.-C. Yoccoz. *Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne*. Ann. Scien. ÉNS **17** (1984) 333–359.
- [Yoc2] J.-C. Yoccoz. *Introduction to hyperbolic dynamics*. In "Real and complex dynamical systems" (Hillerød, 1993), pp 265–291, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **464**, Kluwer, 1995.
- [You] L.-S. Young. *Recurrence times and rates of mixing*. Israel J. Math. **110** (1999) 153–188.
- [Zim] R. J. Zimmer. *Ergodic theory and semisimple groups*. Birkhäuser, 1984.

Frédéric Paulin  
 Laboratoire de mathématique d'Orsay, UMR 8628 CNRS  
 Université Paris-Saclay, 91405 ORSAY Cedex, FRANCE  
 courriel : frederic.paulin@universite-paris-saclay.fr