

Introduction 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 2022-2023, version préliminaire

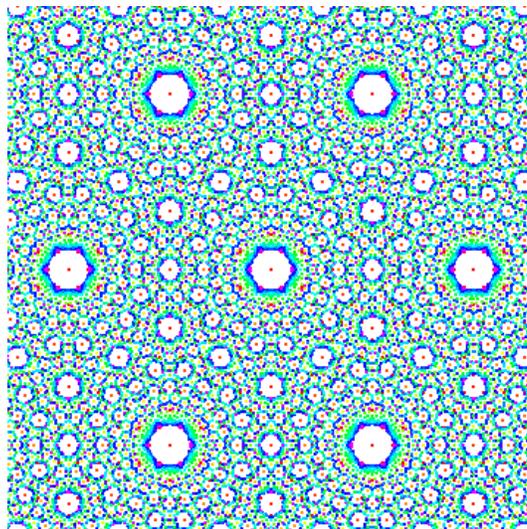


Table des matières

Préambule	4
1 Vocabulaire des systèmes dynamiques et exemples fondamentaux	5
1.1 Orbites périodiques	6
1.2 Parties invariantes et mesures invariantes	12
1.3 Récurrence topologique	18
1.4 L'ensemble (topologiquement) non errant	19
1.5 Transitivité (topologique), mélange topologique et minimalité	20
1.6 Récurrence multiple	25
1.7 Exercices	28
1.8 Indications pour la résolution des exercices	29
2 Dynamique des homéomorphismes du cercle	34
2.1 Nombre de rotation d'un homéomorphisme du cercle	35
2.2 Théorie de Poincaré-Denjoy ⁴³ des homéomorphismes du cercle	38
2.3 Problème de C^r -conjugaison et petits diviseurs	44
2.4 Exercices	45
2.5 Indications pour la résolution des exercices	47
3 Dynamique linéaire des tores	50
3.1 Séries de Fourier sur le tore	50
3.2 Unique ergodicité des translations sur le tore	52
3.3 Propriétés de mélange des endomorphismes linéaires du tore	52
3.4 Exercices	60
3.5 Indications pour la résolution des exercices	63
4 Dynamique sur les espaces homogènes	69
4.1 Le groupe unimodulaire $SL_N(\mathbb{R})$	69
4.2 Représentations unitaires	72
4.3 Ergodicité des translations sur les espaces homogènes	74
4.4 Mélange des translations sur les espaces homogènes	76
4.5 Flots géodésique et horocyclique sur les surfaces hyperboliques	80
4.6 Exercices	90
4.7 Indications pour la résolution des exercices	94
5 Systèmes dynamiques hyperboliques	101
5.1 Endomorphismes linéaires hyperboliques	101
5.2 Le théorème de Grobman-Hartman	108
5.3 Stabilité structurelle des automorphismes hyperboliques du tore	110
5.4 Ensembles hyperboliques de difféomorphismes	114
5.4.1 Fibrés vectoriels normés	114
5.4.2 Parties hyperboliques d'une variété munie d'un difféomorphisme	117
5.5 Le fer à cheval de Smale	124
5.6 Le critère des champs de cônes et la stabilité des ensembles hyperboliques	127
5.7 Expansivité et lemme de pistage des systèmes dynamiques hyperboliques	132
5.8 Le théorème des variétés stables et instables	139

5.8.1	Variétés stables des perturbations d'endomorphismes linéaires	139
5.8.2	Variétés stables et instables locales des difféomorphismes	143
5.8.3	Variétés stables et instables globales des difféomorphismes	147
5.9	La structure de produit local	149
5.10	Exercices	151
5.11	Appendice : rappels de théorie spectrale	154
	Sur la complexification des espaces de Banach réels	155
5.12	Indications pour la résolution des exercices	160
6	Mesures d'entropie maximale	166
6.1	Entropie topologique	166
	Recouvrements ouverts	166
	Définition de l'entropie topologique	168
	Propriétés élémentaires de l'entropie topologique	169
6.2	Rappels sur l'entropie mesurée	171
6.3	Principe variationnel	174
6.4	Entropie topologique des systèmes dynamiques symboliques	179
	6.4.1 Sous-décalages de type fini	180
	6.4.2 Chaînes de Markov sur un alphabet fini	183
	6.4.3 Problématique de l'équidistribution	188
	6.4.4 Mesures d'entropie maximale des sous-décalages de type fini	190
6.5	Entropie topologique des automorphismes linéaires du tore	194
	6.5.1 Partition de Markov du tore	195
	6.5.2 Codage des automorphismes linéaires hyperboliques du tore	199
	6.5.3 Mesure d'entropie maximale des automorphismes hyperboliques du tore	200
6.6	Exercices	203
6.7	Indications pour la résolution des exercices	205
	Index	209
	Références	214

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 ϕ^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 ϕ (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 ϕ . Dans ce dernier cas, on s'intéresse au comportement asymptotique des itérations ϕ^n de ϕ 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 topologie [système dynamique topologique, voire système dynamique différentiable quand l'espace des phases est une variété différentielle] et celui de la théorie de la mesure [théorie ergodique].

Nous étudierons dans ce cours le premier point de vue, et ses interactions avec le second, en renvoyant au cours de seconde année de Master de l'Université Paris-Saclay de Théorie ergodique ou [PY, Wal, CorFS, Pet, BenP, EW, Cou] pour son apprentissage. 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. 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, Nikodim et Stone, venant de <https://mathshistory.st-andrews.ac.uk/Biographies>.

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

1 Vocabulaire des systèmes dynamiques et exemples fondamentaux

Un *système dynamique*¹ est un couple (\widehat{X}, ϕ) (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 (X, \mathcal{B}, μ) (on parle alors de *système dynamique mesuré*), ou
- un espace topologique (X, \mathcal{O}) (on parle alors de *système dynamique topologique*), ou
- une variété différentielle (lisse, de dimension d)² (X, \mathcal{A}) (on parle alors de *système 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 paramètre³, appelé le *flot* du système dynamique à temps continu, qui vérifie respectivement

- ϕ est mesurable ou ϕ^t est mesurable pour tout $t \in \mathbb{R}$,
- ϕ préserve la mesure⁴ ou ϕ^t préserve la mesure pour tout $t \in \mathbb{R}$,
- ϕ est continue ou l'application $(t, x) \mapsto \phi^t(x)$ de $\mathbb{R} \times X$ dans X est continue,
- ϕ est différentiable⁵ 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 tribu⁶, la mesure, la topologie ou l'atlas, nous noterons par abus (X, ϕ) au lieu de (\widehat{X}, ϕ) à partir de maintenant.

Exemples. (1) (**Inversibilité**) Un système dynamique à temps discret (X, ϕ) est dit *inversible* si l'application ϕ est bijective, et si (X, ϕ^{-1}) est encore un système dynamique du même type (mesurable, mesuré, topologique, différentiable de même classe). Le temps

1. 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.

2. c'est-à-dire un espace topologique métrisable séparable X muni d'un atlas de cartes \mathcal{A} de classe C^∞ 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, φ) 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 Γ 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, φ) 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 $\mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$.

3. c'est-à-dire une famille 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 $s, t \in \mathbb{R}$

4. Si (Y, \mathcal{C}, ν) et (Y', \mathcal{C}', ν') sont deux espaces mesurés, une application $f : Y \rightarrow Y'$ préserve la mesure si f est mesurable et $f_* \nu = \nu'$, où $f_* \nu$ est la *mesure image* de ν par f , définie par $f_* \nu(C') = \nu(f^{-1}(C'))$ pour tout $C' \in \mathcal{C}'$. Notons que si $g : Y' \rightarrow Y''$ est une autre application mesurable à valeurs dans un espace mesuré $(Y'', \mathcal{C}'', \nu'')$, alors

$$(g \circ f)_* = g_* \circ f_* .$$

5. c'est-à-dire que ϕ est continue et que pour toutes les cartes (U, φ) et (V, ψ) 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. aussi appelée σ -algèbre

1 d'un système dynamique à temps continu $(X, (\phi^t)_{t \in \mathbb{R}})$ est le système dynamique à temps discret (X, ϕ^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}, ϕ) est un système dynamique mesurable. Si (X, \mathcal{O}, ϕ) est un système dynamique topologique et si \mathcal{B} est la tribu des boréliens de X , alors (X, \mathcal{B}, ϕ) est un système dynamique mesurable. Si (X, \mathcal{A}, ϕ) 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}, ϕ) est un système dynamique topologique. Pour tous les entiers $k \leq k'$, si (X, ϕ) est un système dynamique différentiable de classe $C^{k'}$, alors (X, ϕ) 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, ϕ) et (X', ϕ') deux systèmes dynamiques, tous deux à temps discret (respectivement à temps continu). Une *semi-conjugaison ensembliste* de (X, ϕ) dans (X', ϕ') 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, ϕ) et (X', ϕ') clairement identifiés comme mesurables, mesurés, topologiques ou différentiables de classe C^k , une *semi-conjugaison* de (X, ϕ) dans (X', ϕ') est une semi-conjugaison ensembliste qui est de plus une application mesurable, préservant la mesure⁷, continue ou différentiable de classe C^k respectivement. Le système dynamique (X', ϕ') 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 inverse⁸, 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.

1.1 Orbites périodiques

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

Si (X, ϕ) 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

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

7. Dans le cas de deux systèmes dynamiques mesurés, on omet la condition de surjectivité de la semi-conjugaison, car la condition de préserver la mesure implique que l'image est de mesure totale.

8. On demande seulement le fait que la conjugaison et son inverse soient définis sur un ensemble de mesure totale.

(où $\phi^0 = \text{id}$ et par récurrence⁹ $\phi^{n+1} = \phi \circ \phi^n = \phi^n \circ \phi$ pour tout $n \in \mathbb{N}$). Si (X, ϕ) 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, ϕ) 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, ϕ) 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

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

Si (X, ϕ) 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, ϕ) 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, ϕ) .

La *période* d'un point périodique x d'un système dynamique à temps discret (X, ϕ) est le minimum des $n \in \mathbb{N} - \{0\}$ tels que $\phi^n(x) = x$, ou de manière équivalente, le cardinal de 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 inférieure¹⁰ des $t > 0$ tels que $\phi^t(x) = x$.

Remarques. (i) Si $h : X \rightarrow X'$ est une semi-conjugaison (ensembliste) entre deux systèmes dynamiques (X, ϕ) et (X', ϕ') 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, ϕ^t) et (X', ϕ'^t) . La semi-conjugaison h envoie point périodique sur point périodique :

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

envoie toute orbite positive de (X, ϕ) dans une orbite positive de (X', ϕ') :

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

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

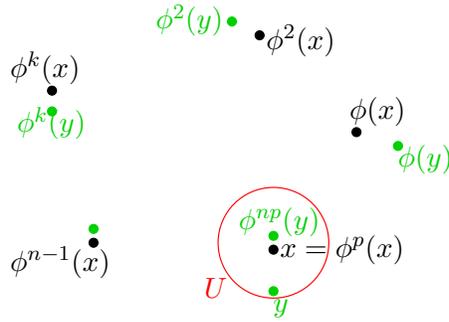
$$\forall x \in X, \quad h(\mathcal{O}(x)) \subset \mathcal{O}(h(x)) . \tag{3}$$

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

(ii) Soient (X, ϕ) 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 $\lim_{n \rightarrow +\infty} \phi^{pn}(y) = x$.

9. 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 $\phi(x)^n = \phi(x) \times \dots \times \phi(x)$.

10. 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.



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, ϕ) est inversible, nous dirons que le point périodique x est *répulsif* s'il est attractif pour le système dynamique (X, ϕ^{-1}) . Si h est une conjugaison entre (X, ϕ) et un système dynamique topologique à temps discret (X', ϕ') , alors x est un point périodique attractif (respectivement répulsif) de (X, ϕ) si et seulement si $h(x)$ est un point périodique attractif (respectivement répulsif) de (X', ϕ') . Nous renvoyons au troisième exemple ci-dessous pour une illustration.

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 θ . Une rotation irrationnelle n'a pas d'orbite périodique, et toute orbite est dense. ¹¹

(2) Soit \mathcal{A} un ensemble non vide, muni de la topologie discrète, appelé un *alphabet*. ¹² Le *système de Bernoulli* ¹³ *unilatère* (respectivement *bilatère*) d'alphabet \mathcal{A} est le système dynamique topologique à temps discret (Σ_+, σ_+) (respectivement (Σ, σ)) où

- 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 σ_+ (respectivement σ), appelée *décalage de Bernoulli à gauche* est l'application qui à une suite $(x_n)_n$ associe la suite $(y_n)_n$ où $y_n = x_{n+1}$ pour tout n dans \mathbb{N} (respectivement \mathbb{Z}).

11. 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 θ .

12. Notons que \mathcal{A} peut être remplacé par n'importe quel espace mesurable (Ω, \mathcal{B}) pour construire deux systèmes dynamiques mesurables $(\Sigma_+ = \Omega^{\mathbb{N}}, \sigma_+)$ et $(\Sigma = \Omega^{\mathbb{Z}}, \sigma)$. Voir le cours de Théorie ergodique ou [PY, Wal, CorFS, Pet, EW, Cou].



13. Jacques Bernoulli
1654-1705



Georg Cantor
1845-1918



Olexandr Sharkovsky
1936-



Stephan Banach
1892-1945

Le système de Bernoulli unilatère (Σ_+, σ_+) est non inversible si le cardinal de \mathcal{A} est au moins 2. Le système de Bernoulli bilatère (Σ, σ) 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 Σ_+ et Σ des systèmes de Bernoulli sont totalement discontinus¹⁴ et métrisables respectivement pour les distances

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

avec les conventions usuelles¹⁵ 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 Σ_+ et Σ sont compacts, homéomorphes à un espace de Cantor^{13, 16}.

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

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

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 (Σ_+, σ_+) est une base d'ouverts de Σ_+ . Appelons *cylindre* (centré) du système de Bernoulli bilatère toute partie de Σ de la forme

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

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 (Σ, σ) est une base d'ouverts de Σ .

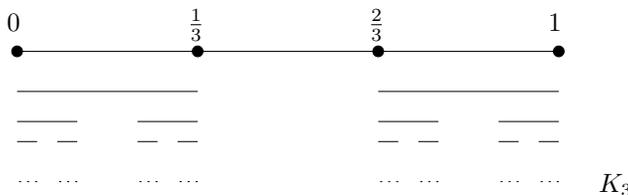
Les points périodiques de σ_+ (respectivement σ) sont les mots $(x_n)_{n \in \mathbb{N}}$ (respectivement $(x_n)_{n \in \mathbb{Z}}$) périodiques, c'est-à-dire tels qu'il existe $k \in \mathbb{N} - \{0\}$ tel que $x_{n+k} = x_n$ pour tout n dans \mathbb{N} (respectivement \mathbb{Z}). Lorsque \mathcal{A} 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 Σ_+ et Σ respectivement. En effet, dans le cas bilatère (l'argument est similaire dans le cas unilatère),

14. 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].

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

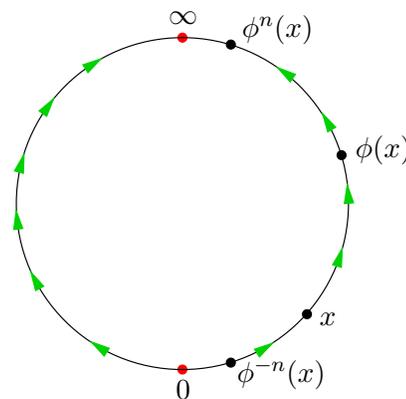
16. Rappelons (voir par exemple [Dug]) qu'un *espace de Cantor* est un espace topologique compact métrisable totalement discontinu sans point isolé, et qu'ils sont tous homéomorphes à *l'espace de Cantor triadique* $K_3 = \{ \sum_{i=1}^{\infty} \frac{a_i}{3^i} : a_i \in \{1, 2\} \}$.



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 .

(3) Notons $X = \widehat{\mathbb{R}}$ le compactifié d'Alexandrov¹⁷ 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 $\infty \mapsto \infty$.

Ce système a exactement deux points périodiques, qui sont en fait des points fixes, les points 0 et ∞ . 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 ∞ 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 ∞ 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, ϕ) 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 1.1. [Théorème de Sharkovsky¹³ (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.

17. 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 ∞ 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].

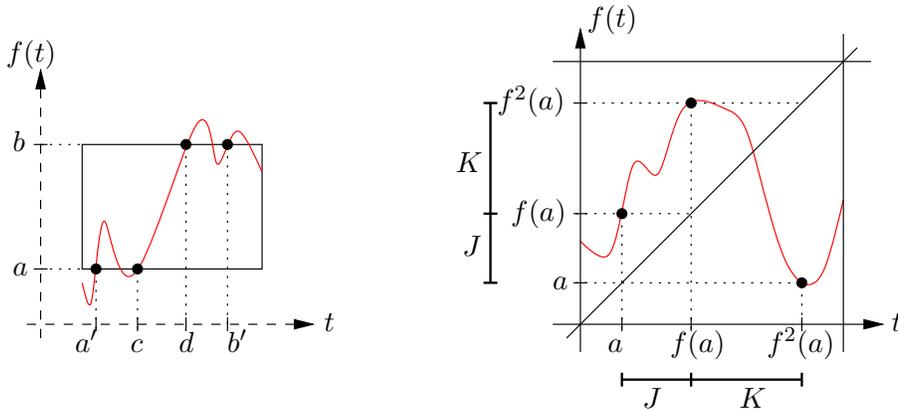
Lemme 1.2. Soient A et B deux intervalles compacts de \mathbb{R} , et $g : A \rightarrow \mathbb{R}$ une application continue.

(1) Si $g(A)$ contient A , alors g admet un point fixe dans A .

(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 1.1, 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 $x \mapsto -x$,¹⁸ 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 1.2 (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 1.2 (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 1.2 (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 $f^{n+2}(L_{n+1})$

18. c'est-à-dire quitte à remplacer f par $g : x \mapsto -f(-x)$

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 1.2 (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 $p \geq 1$ 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

1.2 Parties invariantes et mesures invariantes

Soit (X, ϕ) 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, ϕ) est un système dynamique topologique, alors l'adhérence \bar{Y} de toute partie positivement invariante Y l'est encore.¹⁹

Soient (X, ϕ) 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 . Alors $(Y, \phi|_Y)$ est un système dynamique topologique ou différentiable respectivement, appelé la *restriction* de (X, ϕ) à Y . Si (X, ϕ) est à temps discret et inversible, ou 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 .

Si (X, ϕ) est un système dynamique mesurable (respectivement mesuré) à temps discret, une partie mesurable Y de X est dite *invariante* 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. Si Y est munie de la tribu induite (respectivement de la tribu induite et de la mesure induite), alors $(Y, \phi|_Y)$ est un système dynamique mesurable (respectivement mesuré), appelé la *restriction* de (X, ϕ) à Y .

Si (X, \mathcal{B}, ϕ) est un système dynamique mesurable, et si μ est une mesure sur (X, \mathcal{B}) , rappelons que μ est dite *invariante* si $\phi_*\mu = \mu$ en temps discret, ou $(\phi^t)_*\mu = \mu$ pour tout $t \in \mathbb{R}$ en temps continu. Par exemple, la mesure de comptage sur une orbite périodique d'un système dynamique à temps discret est une mesure invariante.

Si (X, ϕ) est un système dynamique topologique à temps discret, nous noterons

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

19. 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 $\phi(y)$.

l'ensemble des mesures de probabilité boréliennes sur X invariantes par ϕ . 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 signées²⁰ sur X , muni de la topologie vague²¹.

Proposition 1.3. *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 vide²², et $\text{Prob}_{\phi}(X)$ est l'enveloppe convexe fermée²³ de ses points extrémaux.*

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 1.3. 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 Banach¹³ réel des applications continues de X dans \mathbb{R} qui s'annulent à l'infini, muni de la *norme uniforme*

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

20. Une *mesure signée* sur un espace mesurable (X, \mathcal{A}) est une application $\mu : \mathcal{A} \rightarrow \mathbb{R}$ qui est σ -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

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

et telle que $\mu(\emptyset) = 0$. 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 μ 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} . Voir par exemple [Coh, §4].

21. 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 de $|\mu_i|$ est uniformément bornée, il suffit de vérifier ceci pour tout f dans une partie dense de $C^0(X, \mathbb{R})$.

22. Ce résultat d'existence est parfois appelée le *théorème de Krylov-Bogolyubov*.

23. 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 .

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 régulières²⁴ sur X , muni de la norme de la *variation totale*

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

Par le *théorème de représentation de Riesz*²⁵ (voir [Coh, page 220]), l'application de $\mathcal{M}_{\mathbb{R}}^r(X)$ dans le dual topologique²⁶ $C_0^0(X; \mathbb{R})^\vee$ de $C_0^0(X; \mathbb{R})$, définie par

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

est un isomorphisme linéaire isométrique pour la norme de la variation totale sur $\mathcal{M}_{\mathbb{R}}^r(X)$ et la norme duale sur $C_0^0(X; \mathbb{R})^\vee$, par lequel nous identifions désormais ces deux espaces.

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 ℓ 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).

Lemme 1.4. *Étant donné une application continue propre²⁷ $\varphi : Y \rightarrow Z$ entre deux espaces topologiques localement compacts, l'application $\mu \mapsto \varphi_*\mu$ de mesure image par φ , de $\mathcal{M}_{\mathbb{R}}^r(Y)$ dans $\mathcal{M}_{\mathbb{R}}^r(Z)$, est linéaire et continue pour la topologie vague.*

24. 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 32), alors toute mesure de probabilité est régulière, voir [Coh, page 206].



Frigyes Riesz
25. 1880-1956



Mark Krein
1907-1989



Johan Radon
1887-1956



Otto Nikodym
1887-1974

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

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

27. 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.

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 φ 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 1.5. *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-étoile²⁸). 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 1.3. 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 ϕ , par la linéarité de ϕ_* . Donc $\text{Prob}(X)$ est convexe. Par la continuité de ϕ_* (voir le lemme 1.4), 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) ν sur X (par exemple la masse de Dirac Δ_x en un point x de 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 métrisabilité de $\text{Prob}(X)$ (voir la proposition 1.5) assure qu'il existe une sous-suite $(\nu_{n_k})_{k \in \mathbb{N}}$ qui converge vers un élément μ 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 1.4 de continuité de ϕ_* .

La dernière assertion de la proposition 1.3 découle du *théorème de Krein²⁵-Milman* disant qu'un convexe compact d'un espace vectoriel topologique localement convexe séparé est l'enveloppe convexe fermée de ses points extrémaux. \square

La proposition suivante caractérise les points extrémaux du convexe compact non vide $\text{Prob}_{\phi}(X)$, et nous renvoyons au cours de Théorie ergodique ou à [PY, Wal, CorFS, Pet, EW, Cou] pour des développements.

Proposition 1.6. *Soient (X, \mathcal{B}) un espace mesurable, $\phi : X \rightarrow X$ une application mesurable et μ une mesure de probabilité sur (X, \mathcal{B}) invariante par ϕ . Les assertions suivantes sont équivalentes.*

- (1) *La transformation ϕ est ergodique²⁹ pour μ , c'est-à-dire que les seules parties mesurables invariantes sont de mesure nulle ou pleine :*

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

²⁸. Ces deux topologies coïncident de manière évidente si X est compact, car alors nous avons $C_c^0(X; \mathbb{R}) = C_0^0(X; \mathbb{R}) = C^0(X; \mathbb{R})$.

²⁹. Il est possible de définir ainsi l'ergodicité pour des mesures qui ne sont pas de probabilité, et pour des transformations qui ne font que préserver la classe de la mesure, voir par exemple [Aar].

- (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 $\mathbb{L}^2(X, \mu)$, telle que $f \circ \phi = f$ presque partout, est presque partout constante.
- (5) La mesure μ est un point extrémal du convexe $\text{Prob}_\phi(X, \mathcal{B})$ des mesures de probabilité invariantes par ϕ sur (X, \mathcal{B}) , dans l'espace vectoriel $\mathcal{M}_\mathbb{R}(X, \mathcal{B})$ des mesures signées sur (X, \mathcal{B}) .

Démonstration. Nous renvoyons au cours de Théorie ergodique ou à [PY, Wal, CorFS, Pet, EW, Cou] pour l'équivalence des quatre premières assertions. Montrons que μ est ergodique si et seulement si, pour toutes les mesures de probabilité μ_1, μ_2 invariantes par ϕ 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 μ ergodique. Soient $\mu_1, \mu_2 \in \text{Prob}_\phi(X, \mathcal{B})$ et $t \in]0, 1[$ tels que $\mu = t\mu_1 + (1-t)\mu_2$. Les ensembles de mesure nulle pour μ sont aussi de mesure nulle pour μ_1 . Le théorème³⁰ de Radon²⁵-Nikodym²⁵ (voir [Coh, page 132]) assure qu'il existe une fonction f dans $\mathbb{L}^1(X, \mu)$ telle que $\mu_1 = f\mu$. Comme μ et μ_1 sont invariantes par ϕ , nous avons $f = f \circ \phi$ presque partout pour μ . Comme μ est ergodique, l'application f est constante presque partout et les mesures de probabilité μ et μ_1 sont égales. Donc $\mu_1 = \mu_2$ et μ est extrémale.

Réciproquement, supposons que μ 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 $A_2 = {}^cA_1$. Posons alors $\mu_i = \frac{1}{\mu(A_i)}\mu|_{A_i}$ pour $i = 1, 2$. Les mesures de probabilité μ_i sont invariantes par ϕ , différentes de μ et $\mu = \mu(A_1)\mu_1 + (1 - \mu(A_1))\mu_2$. Donc μ n'est pas extrémale. \square

La propriété d'ergodicité d'un système dynamique mesuré est souvent utilisée en conjonction avec le théorème suivant, pour la démonstration duquel nous renvoyons au cours de Théorie ergodique ou à [PY, Wal, CorFS, Pet, EW, Cou]. 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 $\mu(A)$.

Théorème 1.7. (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

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

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

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

30. qui s'applique car μ_1 et μ_2 sont des mesures de probabilité, donc des mesures σ -finies

(iii) $\|\tilde{f}\|_{\mathbb{L}^1} \leq \|f\|_{\mathbb{L}^1}$.

(iv) Si la mesure μ est finie, alors la convergence a lieu dans \mathbb{L}^1 :

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

(v) Pour toute partie A dans \mathcal{B} invariante par ϕ , de mesure finie, nous avons

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

(vi) En particulier, si μ est une mesure de probabilité ergodique, alors

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

pour μ -presque tout $x \in X$. □

Une mesure de probabilité invariante, dont l'existence est assurée par exemple sous les hypothèses de la proposition 1.3, est rarement unique (considérer par exemple ϕ valant l'identité). Voici un critère d'unicité, pour la démonstration duquel nous renvoyons au cours de Théorie ergodique ou à [PY, Wal, CorFS, Pet, EW, Cou].

Proposition 1.8. *Soit 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, ϕ) est uniquement ergodique, c'est-à-dire il admet une unique mesure de probabilité invariante ergodique.*
- (2) *Le système dynamique topologique (X, ϕ) admet une unique mesure de probabilité invariante.*
- (3) *Pour tout $f \in C^0(X, \mathbb{C})$, et pour tout x dans X , les moyennes orbitales*

$$S_n f(x) = \frac{1}{n} \sum_{k=0}^{n-1} f \circ \phi^k(x)$$

convergent vers une limite $L(f)$ qui ne dépend pas de x .

Nous avons alors $L(f) = \mu(f)$, où μ est l'unique mesure de probabilité invariante. □

Nous construirons dans les parties 6.4 et 6.5 de nombreuses mesures invariantes ergodiques pour les systèmes dynamiques symboliques et les endomorphismes linéaires du tore. Nous renvoyons à la proposition 3.2 (1) pour des cas d'unicité.

Exercice E.1. ³¹ Soient X un espace topologique à base dénombrable d'ouverts ³², μ une mesure de probabilité borélienne sur X de support ³³ égal à X , et $\phi : X \rightarrow X$ une application

31. Voir la partie 1.8 pour des indications de solutions.

32. 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.

33. 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.

continue qui préserve μ . Montrer que si ϕ est ergodique, alors μ -presque toute orbite positive du système dynamique topologique à temps discret (X, ϕ) est dense dans X (c'est-à-dire, pour μ -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.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 ϕ_1 et ϕ_2 .

1.3 Récurrence topologique

Soient (X, ϕ) un système dynamique topologique (ou différentiable), et $x \in X$. L'ensemble ω -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, ϕ) est à temps discret et inversible, ou est à temps continu, alors nous définissons de même l'ensemble α -limite de x pour (X, ϕ) , noté $\alpha(x)$ ou $\alpha_\phi(x)$, comme étant l'ensemble ω -limite de x pour la transformation inverse ou le flot inverse.

Le point x est dit *positivement récurrent*³⁴ pour (X, ϕ) si

$$x \in \omega(x).$$

Si (X, ϕ) 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, ϕ) est à temps discret et inversible, ou est à temps continu), et son ensemble ω -limite (et son ensemble α -limite si (X, ϕ) 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 ω -limite est un fermé positivement invariant, qui est non vide si X est compact.

Exercice E.3. Soit (X, ϕ) un système dynamique topologique (ou différentiable), à temps discret et inversible, ou à temps continu. Montrer que tout ensemble ω -limite est globalement invariant.

34. 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** A (où A est une partie mesurable donnée) d'un système dynamique mesurable ou mesuré à temps discret (X, ϕ) , où x est (infiniment positivement) récurrent dans A s'il existe une infinité de $n \in \mathbb{N} - \{0\}$ tels que $\phi^n(x) \in A$.

Exercice E.4. Soit (X', ϕ') un système dynamique topologique (ou différentiable), et soit $h : X \rightarrow X'$ une semi-conjugaison entre (X, ϕ) et (X', ϕ') . Montrer que

$$h(\omega_\phi(x)) \subset \omega_{\phi'}(h(x)), \quad (5)$$

avec égalité si h est une conjugaison.

1.4 L'ensemble (topologiquement) non errant

Soit (X, ϕ) un système dynamique topologique (ou différentiable).

Un point $x \in X$ est dit (topologiquement) *errant* pour (X, ϕ) 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, ϕ) . Il est noté $\Omega(\phi)$ (ou $\Omega(X, \phi)$ lorsqu'il convient de préciser).³⁵ Le système dynamique topologique (X, ϕ) 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, ϕ) est à temps discret et inversible, ou s'il est à temps continu, alors x est errant pour (X, ϕ) si et seulement s'il l'est pour le système dynamique inverse.

Proposition 1.9. *L'ensemble non errant $\Omega(\phi)$ de (X, ϕ) est un fermé positivement invariant, qui est globalement invariant si (X, ϕ) est à temps discret et inversible, ou à temps continu.*

Le support de toute mesure (positive borélienne) finie sur X invariante par ϕ 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, ϕ) 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 μ une mesure sur X invariante par ϕ . 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 μ est finie, nous avons $\mu(U) = 0$ et x n'appartient pas au support de μ . \square

Exemples. (1) Un point périodique ou positivement récurrent est non errant.

(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.

³⁵. 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 $n \in \mathbb{N} - \{0\}$.

Exercice E.5. Soit (X', ϕ') un système dynamique topologique (ou différentiable), et soit $h : X \rightarrow X'$ une semi-conjugaison entre (X, ϕ) et (X', ϕ') . Montrer que

$$h^{-1}(\Omega(\phi')) \supset \Omega(\phi), \quad (6)$$

avec égalité si h est une conjugaison.

1.5 Transitivité (topologique), mélange topologique et minimalité

Soit (X, ϕ) 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, ϕ) 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.7 pour montrer ou utiliser qu'un système dynamique est positivement transitif ou transitif.

Exercice E.6. Soit $h : X \rightarrow X'$ une semi-conjugaison entre (X, ϕ) et un système dynamique topologique (X', ϕ') . Montrer que si (X, ϕ) est positivement transitif (respectivement à temps discret et inversible, ou à temps continu, et transitif), alors le facteur (X', ϕ') 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.10.

Exercice E.7. Soient X un espace de Baire³⁶ à base dénombrable d'ouverts, et $\phi : X \rightarrow X$ une application continue. Montrer que les affirmations suivantes sont équivalentes :

- (1) la transformation ϕ est positivement transitive,
- (2) il existe un point dans X dont l'ensemble ω -limite est égal à X ,
- (3) l'ensemble Y des points dans X dont l'ensemble ω -limite est égal à X est un G_δ dense de X .

Si ϕ est un homéomorphisme, montrer que les affirmations suivantes sont équivalentes :

- (i) la transformation ϕ 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_δ dense de X .

Le système dynamique (X, ϕ) 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

36. Rappelons qu'un *espace de Baire* 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_δ (dense) dans X toute intersection dénombrable d'ouverts denses.

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.6 (voir l'exercice E.9), si $h : X \rightarrow X'$ est une semi-conjugaison entre (X, ϕ) et un système dynamique topologique (X', ϕ') , et si (X, ϕ) est topologiquement mélangeant, alors le facteur (X', ϕ') 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 (Σ_+, σ_+) 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 σ_+ 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 ω -limite de x est égal à tout Σ_+ . Le cas bilatère (Σ, σ) 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 σ 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 Σ_+ .

Le système dynamique (X, ϕ) est dit *positivement minimal* si l'une des conditions équivalentes³⁷ suivantes est vérifiée :

- (1) toute orbite positive est dense,
- (2) l'ensemble ω -limite de tout point de X est égal à X ,
- (3) les seuls fermés positivement invariants de X sont \emptyset et X .

³⁷. 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

$$\omega(x) = \bigcap_{t \geq 0} \overline{\mathcal{O}^+(\phi^t(x))} = \bigcap_{t \geq 0} X = X.$$

Si (X, ϕ) 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, ϕ) 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).

(ii) Une rotation irrationnelle du cercle est minimale, car toute orbite est dense.

Remarques. (1) Supposons (X, ϕ) à temps discret et inversible, ou à temps continu. Si le système dynamique topologique (X, ϕ) 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 ω -limite est fermé, non vide par compacité et globalement invariant par l'exercice E.3, donc égal à X si (X, ϕ) est minimal.

(2) Comme la frontière³⁸ 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 $\partial M = M - \overset{\circ}{M}$.

(3) Tout système dynamique topologique (X, ϕ) , 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 1.1).

Proposition 1.10. *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

38. 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 ∂A est l'ensemble des points de X dont tout voisinage rencontre à la fois A et son complémentaire.

minimal par compacité, voir la remarque (1) ci-dessus. Ceci résulte du lemme de Zorn³⁹ (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 1.11. (Théorème de récurrence de Birkhoff³⁹) *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 1.11 qui n'utilise pas l'axiome du choix, sous l'hypothèse supplémentaire que X est métrisable. Elle utilise le théorème de récurrence de Poincaré³⁹ ci-dessous, pour une démonstration⁴⁰ duquel nous renvoyons au cours de Théorie ergodique ou à [PY, Wal, CorFS, Pet, EW, Cou].

Théorème 1.12. (Théorème de récurrence de Poincaré) *Soient (X, \mathcal{B}, μ) un espace de probabilité, $\phi : X \rightarrow X$ une application mesurable préservant μ , et A un élément de \mathcal{B} . Alors, pour μ -presque tout point x de A , l'orbite $(\phi^n(x))_{n \in \mathbb{N}}$ repasse une infinité de fois dans A .* \square

Montrons que le théorème 1.12 implique le théorème 1.11 sous l'hypothèse supplémentaire à celles du théorème 1.11 que X est métrisable. 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 1.3, il existe une mesure de probabilité μ invariante par ϕ . 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 1.12 appliqué à (X, \mathcal{B}, μ) 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 $m \in \mathbb{N}$



Max Zorn
39. 1906–1993



Georges Birkhoff
1884–1944



Henri Poincaré
1854–1912



Henri Lebesgue
1875–1941

40. Voici toutefois une démonstration par souci de complétude.

Démonstration du théorème 1.12. Notons $A_n = \bigcup_{k \geq n} \phi^{-k}(A)$ l'ensemble des points de X qui passent dans A lorsque l'on itère ϕ au moins n fois. Notons $A_\infty = \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} \phi^{-k}(A)$ 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. \square

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.

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 μ invariante.

Supposons de plus ϕ 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, ϕ) est de mesure totale pour toute mesure de probabilité μ invariante.

Exercice E.8. 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'application⁴¹ continue définie par $x \mapsto px$.

(1) Montrer que le système dynamique topologique à temps discret (\mathbb{T}^1, ϕ_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 ϕ_p est dense dans le cercle \mathbb{T}^1 , que le système dynamique topologique (\mathbb{T}^1, ϕ_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 Θ est une semi-conjugaison entre le système de Bernoulli unilatère d'alphabet $\{0, \dots, p-1\}$ et le système dynamique topologique (\mathbb{T}^1, ϕ_p) , que l'on appelle un codage de (\mathbb{T}^1, ϕ_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 λ la mesure de Lebesgue^{39 42} de \mathbb{T}^1 , l'application ϕ_p préserve la mesure de Lebesgue λ , et que nous avons, pour tous les $f, g \in C^0(\mathbb{T}^1)$,

$$\int_{\mathbb{T}^1} (f \circ \phi_p) g \, d\lambda = \int_{\mathbb{T}^1} f(Lg) \, d\lambda. \quad (7)$$

41. appelée l'application de doublement de l'angle si $p = 2$

42. Par définition (voir aussi le début du chapitre 3), 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

$$\int_{\mathbb{T}^1} f \, d\lambda = \int_{t_0}^{t_0+1} f(t) \, dt.$$

Montrer que pour tous les $f \in C^1(\mathbb{T}^1)$ et $n \in \mathbb{N}$, nous avons

$$\left| L^n f(x) - \int_0^1 f(t) dt \right| \leq \frac{1}{p^n} \int_0^1 |f'(t)| dt . \quad (8)$$

En déduire que la transformation ϕ_p est mélangeante pour la mesure de Lebesgue du cercle, c'est-à-dire (voir aussi la partie 3.3) 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 ϕ_p est exponentiellement mélangeante pour la mesure de Lebesgue du cercle (voir la partie 3.3), et plus précisément que pour tous les $f, g \in C^1(\mathbb{T}^1)$, en définissant la norme de Sobolev $W^{1,\infty}$ de f par $\|f\|_{1,\infty} = \max\{\|f\|_\infty, \|f'\|_\infty\}$, nous avons

$$\left| \int_{\mathbb{T}^1} (f \circ \phi_p^n) g d\lambda - \left(\int_{\mathbb{T}^1} f d\lambda \right) \left(\int_{\mathbb{T}^1} g d\lambda \right) \right| \leq \frac{1}{p^n} \|f\|_{1,\infty} \|g\|_{1,\infty} . \quad (9)$$

1.6 Récurrence multiple

Voici une généralisation du théorème de récurrence de Birkhoff 1.11.

Théorème 1.13. (Furstenberg⁴³-Weiss⁴³) Soient X un espace topologique métrisable compact non vide et $f : X \rightarrow X$ une application continue. Alors, pour tout ℓ 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

$$\forall p \in \{1, \dots, \ell\}, \quad \lim_{k \rightarrow \infty} f^{p n_k}(x) = x .$$

Autrement dit, il existe un point x positivement récurrent pour f, f^2, \dots, f^ℓ , avec les mêmes temps de presque-retour. Nous utiliserons le lemme indépendant suivant.

Lemme 1.14. Soient Y un espace topologique métrisable complet et $\varphi : Y \rightarrow \mathbb{R}$ une fonction semi-continue supérieurement⁴⁴. Alors l'ensemble des points de continuité de φ est dense dans Y .

Démonstration. Quitte à remplacer φ par e^φ , nous pouvons supposer que φ est minorée. Notons, pour tout $\epsilon > 0$,

$$Y_\epsilon = \{y \in Y : \exists (y_k)_{k \in \mathbb{N}} \in Y^{\mathbb{N}}, \lim_{k \rightarrow \infty} y_k = y \text{ et } \limsup_{k \rightarrow \infty} \varphi(y_k) \leq \varphi(y) - \epsilon\} .$$



Hillel Furstenberg
1935-



Benjamin Weiss
1941-



Bartel
van der Waerden
1903-1996



Arnaud Denjoy
1884-1974

43.

44. 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.

Comme φ est semi-continue supérieurement, l'ensemble Y_ϵ est fermé⁴⁵ et l'ensemble des points de continuité de φ 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_ϵ est d'intérieur vide.

Si, par l'absurde, l'intérieur de Y_ϵ 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_ϵ telle que $\varphi(y_k) \leq \varphi(y_{k-1}) - \frac{\epsilon}{2}$. Ceci contredirait le fait que φ est minorée. \square

Démonstration du théorème de Furstenberg-Weiss. Par la proposition 1.10, nous pouvons supposer que X est minimal. Procédons par récurrence sur ℓ . Pour $\ell = 1$, c'est le théorème de récurrence de Birkhoff 1.11. 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

$$\forall p \in \{1, \dots, \ell - 1\}, \quad \lim_{k \rightarrow \infty} f^{p n_k}(x) = x.$$

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

$$\begin{aligned} g(x_1, \dots, x_\ell) &= (f(x_1), f^2(x_2), \dots, f^\ell(x_\ell)) \\ h(x_1, \dots, x_\ell) &= (f(x_1), f(x_2), \dots, f(x_\ell)). \end{aligned}$$

Nous voulons trouver un point positivement récurrent de g qui soit sur Δ . 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 Δ 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

$$g^{n_k}(z) = h^m g^{n_k}(z_{0,k}) \xrightarrow{k \rightarrow \infty} h^m(y_0). \quad \square$$

45. En effet, soit $(y^i)_{i \in \mathbb{N}}$ une suite dans Y_ϵ qui converge vers y^∞ 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 φ , nous avons

$$\limsup_{i \rightarrow \infty} \varphi(y_{k_i}^i) \leq \limsup_{i \rightarrow \infty} \varphi(y^i) - \epsilon \leq \varphi(y^\infty) - \epsilon.$$

Étape 3. $\forall \epsilon > 0, \exists y \in \Delta, \exists n \geq 1, d(g^n(y), y) < \epsilon.$

Démonstration. Soient $\epsilon > 0, \epsilon_0 = \frac{\epsilon}{2}$ et y_0 un point quelconque de Δ . L'étape 2 permet de trouver un point y_1 dans Δ 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

$$\forall y \in \Delta, d(y, y_1) < \epsilon_1 \Rightarrow d(g^{n_1}(y), y_0) < \epsilon_0.$$

Recommençons avec ϵ_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 Δ et des réels $\epsilon_k < \epsilon_{k-1}$ tels que

$$\forall y \in \Delta, d(y, y_k) < \epsilon_k \Rightarrow d(g^{n_k}(y), y_{k-1}) < \epsilon_{k-1}.$$

Nous avons alors

$$\forall i < j, d(g^{n_j+n_{j-1}+\dots+n_{i+1}}(y_j), y_i) < \epsilon_i < \frac{\epsilon}{2}.$$

Par la compacité de Δ , 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 φ 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 1.14 ci-dessus, il existe un point de continuité a de φ . 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 Δ 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 Δ pour h . Par compacité de Δ , 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 Δ , il existe $\epsilon > 0$ tel que

$$\forall y', y'' \in \Delta, d(y', y'') < \epsilon \Rightarrow \forall m \in F, d(h^m(y'), h^m(y'')) < \delta.$$

Considérons alors le point y de Δ 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 1.15. (van der Waerden⁴³) 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 ω par le décalage σ_+ , qui est positivement invariante par σ_+ . D'après le théorème de Furstenberg-Weiss, il existe un point ω' 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 ω' est dans l'adhérence de l'orbite de ω , 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

1.7 Exercices

Exercice E.9. Soient (X, ϕ) et (X', ϕ') deux systèmes dynamiques topologiques, et soit $h : X \rightarrow X'$ une semi-conjugaison entre (X, ϕ) et (X', ϕ') . Parmi les propriétés suivantes, quelles sont celles qui sont vérifiées par (X', ϕ') si elles le sont par (X, ϕ) ? 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.10. Soit X un espace topologique sans point isolé, et soit $\phi : X \rightarrow X$ un homéomorphisme transitif.

- (1) Montrer que le système dynamique topologique (X, ϕ) 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 ϕ est positivement transitif.

1.8 Indications pour la résolution des exercices

Correction de l'exercice E.1. 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 μ est total. Par ergodicité de ϕ , 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.2. Soit ν une mesure de probabilité sur X (par exemple la masse de Dirac Δ_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 1.5) assure qu'il existe une sous-suite $(\nu_{n_k})_{k \in \mathbb{N}}$ qui converge vers un élément μ 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 1.4. L'invariance de μ par ϕ_2 se démontre de même.

Correction de l'exercice E.3. Supposons (X, ϕ) 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 ϕ^n est continue et envoie orbite positive sur orbite positive, nous avons⁴⁶

$$\phi^n(y) \in \phi^n(\overline{\mathcal{O}^+(\phi^{m-n}(x))}) \subset \overline{\mathcal{O}^+(\phi^n(\phi^{m-n}(x)))} = \overline{\mathcal{O}^+(\phi^m(x))}.$$

Ceci étant vrai pour tout m assez grand, nous avons $\phi^n(y) \in \omega(x)$, comme souhaité.

Correction de l'exercice E.4. Supposons (X, ϕ) à temps discret (la démonstration est analogue dans le cas à temps continu). Nous avons, par la continuité de h pour la deuxième inclusion,

$$\begin{aligned} h(\omega_\phi(x)) &= h\left(\bigcap_{N \geq 0} \overline{\{\phi^n(x) : n \geq N\}}\right) \subset \bigcap_{N \geq 0} h\left(\overline{\{\phi^n(x) : n \geq N\}}\right) \\ &\subset \bigcap_{N \geq 0} \overline{h(\{\phi^n(x) : n \geq N\})} = \bigcap_{N \geq 0} \overline{\{(\phi')^n(h(x)) : n \geq N\}} \\ &= \omega_{\phi'}(h(x)). \end{aligned}$$

⁴⁶. Rappelons que si $g : Y \rightarrow Z$ est une application continue entre deux espaces topologiques, et si Y' est une partie de Y , alors $g(\overline{Y'}) \subset \overline{g(Y')}$.

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.5. 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 $x \in {}^c\Omega(\phi)$.

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.6. Supposons (X, ϕ) 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', ϕ') est positivement transitif (respectivement transitif).

Correction de l'exercice E.7. 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 ϕ 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_{δ} 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.8. (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 ϕ_p , car

$$p^n \frac{k}{p^n - 1} = \frac{(p^n - 1)k}{p^n - 1} + \frac{k}{p^n - 1} = \frac{k}{p^n - 1} \bmod \mathbb{Z}.$$

Puisque l'ensemble non errant $\Omega(\phi_p)$ de ϕ_p est fermé et contient les points périodiques de ϕ_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

$$\bigcup_{k \in \mathbb{N}} \phi_p^{-k}(x) = \left\{ \frac{x + k}{p^n} \bmod \mathbb{Z} : n \in \mathbb{N}, k \in \mathbb{Z} \right\},$$

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, ϕ_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 ϕ_p^n est

$$\phi_p^{-n}(x) = \left\{ \frac{x + k}{p^n} \bmod \mathbb{Z} : k \in \mathbb{Z} \right\}$$

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 ϕ_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,⁴⁷ et elle semi-conjuge bien le décalage σ_+ et ϕ_p . La surjectivité de Θ vient de l'existence du développement p -adique de tout élément de $[0, 1[$. Puisque Θ est continue et envoie point périodique sur point périodique, et puisque les points périodiques d'un système de Bernoulli sont denses, nous avons

$$\mathbb{R}/\mathbb{Z} = \Theta(\{0, \dots, p-1\}^{\mathbb{N}}) = \Theta(\overline{\text{Per}(\sigma_+)}) \subset \overline{\Theta(\text{Per}(\sigma_+))} \subset \overline{\text{Per}(\phi_p)} \subset \mathbb{R}/\mathbb{Z},$$

⁴⁷ 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.

donc $\text{Per}(\phi_p)$ est dense.

Puisque Θ envoie l'ensemble ω -limite de tout $x \in \Sigma_+$ dans l'ensemble ω -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.7), nous avons donc $\omega_{\phi_p}(\Theta(x)) = \mathbb{R}/\mathbb{Z}$. Donc ϕ_p est positivement transitif.

Puisqu'être topologiquement mélangeant est préservé par passage à un facteur (voir l'exercice E.9) et puisqu'un décalage de Bernoulli est topologiquement mélangeant (voir l'exemple (2) suivant l'exercice E.7), ceci remontre que ϕ_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 $f \in C^0(\mathbb{T}^1)$.

Montrons la formule (7). 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 ϕ_p préserve la mesure de Lebesgue.

Montrons la formule (8). 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 sera vu en toute généralité dans la partie 3.3. Montrons donc la formule (9).

Soient $f, g \in C^1(\mathbb{T}^1)$. Par les formules (7) et (8), 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.9. • 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 ϕ , alors $x \in \omega(x)$. Donc par la formule (5), nous avons

$$h(x) \in h(\omega(x)) \subset \omega(h(x)) ,$$

d'où $h(x)$ est positivement récurrent.

• Nous savons par la formule (6) que $h^{-1}(\Omega(\phi'))$ contient $\Omega(\phi)$. Donc si ϕ n'a pas de point errant, alors, puisque h est surjective, $\Omega(\phi') \supset h(\Omega(\phi)) = h(X) = X'$. D'où ϕ' n'a pas de point errant.

• Si (X, ϕ) est positivement transitif (respectivement transitif dans le cas inversible), alors (X', ϕ') l'est, c'est l'exercice E.6. Voici un autre argument lorsque X et X' sont des espaces de Baire à base dénombrable d'ouverts. Si (X, ϕ) est positivement transitif, soit x un point de X dont l'ensemble ω -limite est égal à tout X . Par la formule (5), l'ensemble ω -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, ϕ) 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', ϕ') est topologiquement mélangeant.

• Supposons que (X, ϕ) soit positivement minimal. Pour tout $x' \in X'$, puisque h est surjective, soit $x \in X$ tel que $h(x) = x'$. Par la formule (5), l'ensemble ω -limite $\omega(x') = \omega(h(x))$ contient $h(\omega(x)) = h(X) = X'$, ces dernières égalités étant vérifiées par la minimalité positive de ϕ et la surjectivité de h . Donc (X, ϕ) est positivement minimal.

De même, si (X, ϕ) 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', ϕ') est minimal.

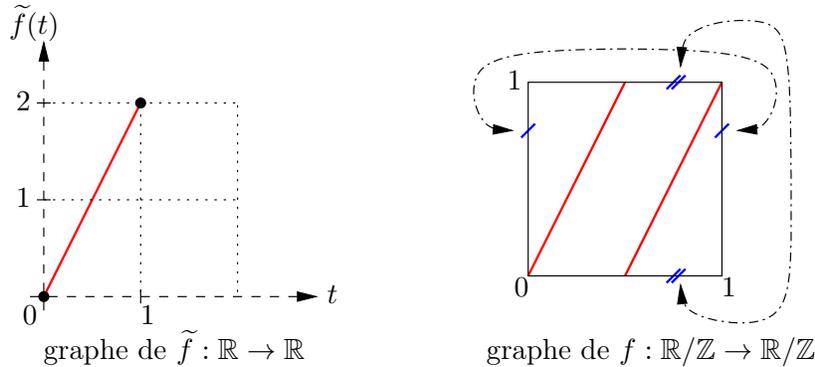
2 Dynamique des homéomorphismes du cercle

Rappelons (voir par exemple [Pau4, Pau8] pour des rappels de topologie algébrique élémentaire) que l'application lisse ρ 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^∞ -difféomorphisme $\mathbb{R}/\mathbb{Z} \rightarrow \mathbb{S}_1$ par lequel nous identifions ces deux variétés différentielles lisses.⁴⁸ Alors $\rho : \mathbb{R} \rightarrow \mathbb{R}/\mathbb{Z}$ est la projection canonique $t \mapsto \dot{t} = t \bmod \mathbb{Z}$. Par le théorème du 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 $x \in \mathbb{R}$.

Exemple. Le degré d'un revêtement de degré d du cercle est égal à d . En particulier, l'application de doublement de l'angle,⁴⁹ qui est l'application $f : x \mapsto 2x$ de \mathbb{R}/\mathbb{Z} dans \mathbb{R}/\mathbb{Z} , est de degré 2.



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 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 α est bien défini modulo \mathbb{Z} , et cette rotation, notée R_α 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

$$R_\alpha : x \mapsto x + \alpha .$$

48. En munissant \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 \mathbb{S}_1 de sa structure de groupe multiplicatif, ce C^∞ -difféomorphisme $\mathbb{R}/\mathbb{Z} \rightarrow \mathbb{S}_1$ est aussi un isomorphisme de groupes.

49. Voir les exercices E.8 et E.12 pour des compléments.

2.1 Nombre de rotation d'un homéomorphisme du cercle

Proposition 2.1. 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

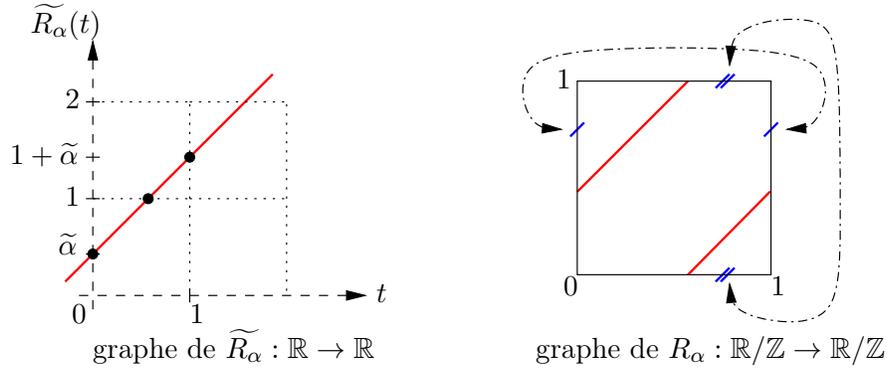
$$\sup_{m \in \mathbb{N}} |\tilde{f}^m(x) - m \tau(\tilde{f}) - x| \leq 2. \quad (10)$$

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 α 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 α dans \mathbb{R} , l'application $\tilde{R}_\alpha : x \mapsto x + \tilde{\alpha}$ est un relevé de R_α (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 2.2. 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 ℓ .

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 $\tilde{f}^m(x+k) = \tilde{f}^m(x) + k$ pour tout $k \in \mathbb{Z}$, que $|\tilde{f}^m(y) - \tilde{f}^m(z)| \leq 1$ 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 2.2 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 (10), en appliquant le lemme 2.2 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 2.3. *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.11. Soit $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$ un homéomorphisme préservant l'orientation, de nombre de rotation α 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_α .

(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 $t \mapsto t + p$.

2.2 Théorie de Poincaré-Denjoy⁴³ des homéomorphismes du cercle

Théorème 2.4 (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 α irrationnel. Alors f est topologiquement conjugué à la rotation R_α .

Remarques. (1) Nous verrons que f n'est pas forcément C^2 -conjugué à R_α . Le théorème reste vrai si f est de classe C^1 et si l'application⁵⁰ $\ln f' : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}$ est à variation bornée⁵¹ (voir [KH, §12.1]).

(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_α . 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

⁵⁰. 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.

⁵¹. 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 $\text{Var}(g) = \frac{1}{2\pi} \int_0^{2\pi} |g'(e^{i\theta})| d\theta$.

coïncident donc sur la partie $\{\tilde{f}^n(0) + m : m, n \in \mathbb{Z}\}$ de \mathbb{R} , qui est dense par conjugaison et car $\{n\alpha + m : n, m \in \mathbb{Z}\}$ l'est (par l'irrationalité de α). Ils coïncident donc partout par continuité de h et h^* .

La première étape de la démonstration du théorème 2.4, due à Poincaré, est de montrer que f est topologiquement semi-conjuguée à R_α , 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 2.5 (Poincaré). *Soit $f : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$ un homéomorphisme préservant l'orientation du cercle, de nombre de rotation α irrationnel. Il existe une application continue monotone⁵² surjective $h : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$ telle que $h \circ f = R_\alpha \circ h$. Si f est transitif,⁵³ alors h est un homéomorphisme, qui conjugue f et R_α .*

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 (10) 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 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 α 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 » de \tilde{h}). Notons U l'image dans \mathbb{R}/\mathbb{Z} de la réunion des intérieurs de ces intervalles, et ω_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.7, 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_α . En particulier, toute orbite de f est dense, et le système dynamique topologique $(\mathbb{R}/\mathbb{Z}, f)$ est minimal.

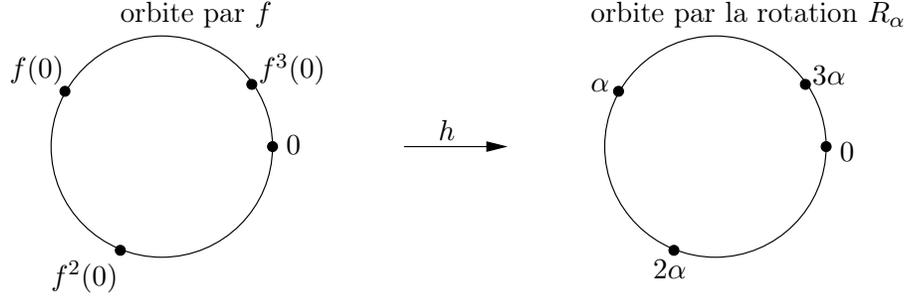
Si f n'est pas transitive, alors ω_f est un espace de Cantor (fermé, sans point isolé et d'intérieur vide du cercle). La préimage d'un point de \mathbb{R}/\mathbb{Z} par la restriction à ω_f de \tilde{h} est constituée de 1 ou 2 points, ces deux points étant alors les extrémités d'une composante

52. 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.

53. c'est-à-dire si le système dynamique topologique $(\mathbb{R}/\mathbb{Z}, f)$ est (topologiquement) transitif

connexe de U . Les points de U sont errants pour f , et l'ensemble non errant de f est ω_f . Toute orbite de ω_f est dense dans ω_f , qui est donc l'unique minimal de f , appelé un *minimal exceptionnel*.

Démonstration 2. La remarque clef pour une deuxième démonstration du théorème 2.5 (donnée par la première assertion du lemme suivant) est que les orbites sur le cercle de l'homéomorphisme f sont ordonnées exactement comme celles de la rotation R_α .



Soit f comme dans l'énoncé du théorème 2.5. 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 2.6. *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 $\tilde{f}^n(x) \neq \tilde{f}^m(x)$ et si I est l'un des deux arcs de cercles d'extrémités $\tilde{f}^n(x)$ et $\tilde{f}^m(x)$, alors toute orbite positive de f rencontre I .*
- (3) *L'ensemble ω -limite d'un point de \mathbb{R}/\mathbb{Z} est indépendant de ce point, et noté ω_f , et ou bien $\omega_f = \mathbb{R}/\mathbb{Z}$ (auquel cas f est transitif), ou bien le fermé ω_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 2.3 (2).

Supposons que

$$\tilde{f}^n(0) + m < \tilde{f}^{n'}(0) + m' . \tag{11}$$

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 (11) 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 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 2.3 (2).

(3) Soient $x, y \in \mathbb{R}/\mathbb{Z}$, montrons que les ensembles ω -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 1.10, le système dynamique topologique (X, f) admet au moins un minimal. Puisque tout minimal de f contient l'ensemble ω -limite de chacun de ses points, le fermé ω_f , qui est non vide par compacité et invariant par f d'après l'exercice E.3, 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 1.5).

Pour tout x dans ω_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 ω_f , alors $f^{n_k}(x) = x$ pour k assez grand, donc x est périodique, ce qui contredit la proposition 2.3 (2). \square

Revenons à la démonstration du théorème 2.5. 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_α 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 $\tilde{A} = \{\tilde{f}^n(0) + m : m, n \in \mathbb{Z}\}$ le relevé total de l'orbite de $0 \in \mathbb{R}/\mathbb{Z}$. Nous définissons $\tilde{h} : \tilde{A} \rightarrow \mathbb{R}$ comme l'unique application strictement croissante envoyant $\tilde{f}^n(0) + m$ sur $n\tilde{\alpha} + m$, ce qui est possible par le lemme 2.6 (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 \tilde{A} ,

$$\tilde{h} \circ \tilde{f} = \tilde{R}_\alpha \circ \tilde{h}.$$

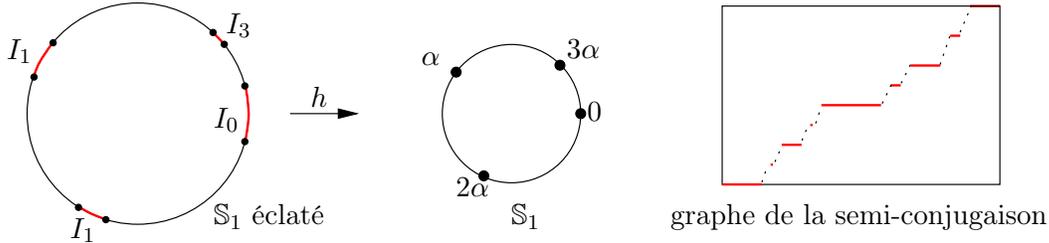
Montrons que \tilde{h} s'étend continûment à l'adhérence $\overline{\tilde{A}}$ de \tilde{A} dans \mathbb{R} . Puisque \tilde{h} est croissante, les limites à droites et à gauche de \tilde{h} existent en tout point de $\overline{\tilde{A}}$, et si elles étaient différentes, le complémentaire de l'adhérence de l'ensemble $\tilde{h}(\tilde{A}) = \{n\tilde{\alpha} + m : m, n \in \mathbb{Z}\}$ serait non vide, ce qui contredirait le fait qu'il est dense dans \mathbb{R} , puisque $\tilde{\alpha}$ est irrationnel. Par densité de $\tilde{h}(\tilde{A})$, l'application $\tilde{h} : \overline{\tilde{A}} \rightarrow \mathbb{R}$ est de plus surjective.

Puisque $\tilde{h} : \overline{A} \rightarrow \mathbb{R}$ est croissante et surjective, pour toute composante connexe I de $\mathbb{R} - \overline{A}$ (qui est un intervalle ouvert et borné par invariance de \overline{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 \overline{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-conjugué f et R_α .

Lorsque f est transitive, nous avons $\overline{A} = \mathbb{R}$ 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_α , et éclatons une ⁵⁴ orbite, au sens suivant. Enlevons du cercle l'orbite complète de $0 \in \mathbb{R}/\mathbb{Z}$ par R_α , et remplaçons le point $R_\alpha^n(0)$ par un intervalle I_n de longueur ℓ_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_α 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 α . L'unique minimal de f est l'ensemble de Cantor adhérence du complémentaire de $\bigcup_{n \in \mathbb{Z}} I_n$. La semi-conjugaison h entre f et R_α 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 2.7 (Denjoy). *Pour tous les $\rho \in (\mathbb{R} - \mathbb{Q})/\mathbb{Z}$ et $\alpha \in]0, 1[$, il existe un*

54. Il serait aussi possible d'éclater un ensemble dénombrable d'orbites.

C^1 -difféomorphisme non transitif du cercle, de dérivée α -höldérienne⁵⁵, de nombre de rotation ρ . \square

Par contre, le but du théorème de Denjoy 2.4 est de montrer que de tels exemples ne sont pas possibles en classe C^2 .

Démonstration du théorème 2.4. 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 2.5.

Supposons par l'absurde que f ne soit pas topologiquement conjugué à la rotation R_α . Alors par le théorème 2.5, l'unique minimal ω_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

$$V = \frac{1}{2\pi} \int_0^{2\pi} |(\ln f')'(e^{i\theta})| d\theta = \sup_{(I'_k)_{1 \leq k \leq n}} \sum_{k=1}^n |\ln f'(y_k) - \ln f'(x_k)|$$

la variation totale de $\ln f'$, où la borne supérieure est prise 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.

Lemme 2.8. *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 2.8. Puisque I s'envoie sur un point par la semi-conjugaison h entre f et la rotation R_α , et puisque l'orbite de I est ordonnée sur le cercle comme celle de

55. Pour tout $\alpha \in]0, 1]$, si X et Y sont des espaces métriques, une application continue $f : X \rightarrow Y$ est dite α -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 α -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.

la rotation irrationnelle R_α 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 $y = f^{-n}(x)$, nous avons

$$\begin{aligned} V &\geq \sum_{k=0}^{n-1} \left| \ln f'(f^k(x)) - \ln f'(f^k(y)) \right| \geq \left| \sum_{k=0}^{n-1} \ln f'(f^k(x)) - \ln f'(f^k(y)) \right| \\ &= \left| \ln \frac{\prod_{k=0}^{n-1} f'(f^k(x))}{\prod_{k=0}^{n-1} f'(f^k(y))} \right| = \left| \ln \frac{(f^n)'(x)}{(f^n)'(y)} \right|, \end{aligned}$$

ce qui montre le lemme 2.8 car $y = f^{-n}(x)$. □

Ceci conclut la démonstration du théorème 2.4. □

2.3 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 $n \in \mathbb{N}$).⁵⁶ Le problème de la régularité d'une conjugaison h entre un C^r -difféomorphisme f du cercle de nombre de rotation α irrationnel et la rotation R_α d'angle α (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,⁵⁷ Moser⁵⁷, Herman⁵⁷ et Yoccoz⁵⁷ entre autres. Je recommande fortement de regarder une présentation historique en français sur YouTube

<https://www.youtube.com/watch?v=9Cfq80fjDw>

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 2.4.

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 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)).

⁵⁶. La régularité C^ω 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.



57. Vladimir Arnold
1937-2010



Jürgen Moser
1929-1999



Michaël Herman
1942-2000



Jean-Christophe Yoccoz
1958-2016



Klaus Roth
1925-2015

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 β* 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 α est dit *diophantien* s'il vérifie une condition diophantienne d'ordre β pour un certain $\beta \geq 0$. Il est dit *de Liouville* s'il n'est pas diophantien. Notons C_β l'ensemble des irrationnels qui vérifient une condition diophantienne d'ordre β , et Dio l'ensemble des nombres diophantiens. Un théorème de Roth⁵⁷ (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 α d'être diophantien est nécessaire à l'existence d'une conjugaison C^∞ pour tout difféomorphisme de nombre de rotation α . Le résultat suivant de Yoccoz, pour lequel nous renvoyons à [Yoc1], dit en particulier que l'ensemble des irrationnels α tels que tout difféomorphisme C^∞ préservant l'orientation du cercle de nombre de rotation α est C^∞ -conjugué à la rotation R_α est exactement l'ensemble Dio des irrationnels diophantiens. En régularité moindre que C^∞ , 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 α -höldérienne.

Théorème 2.9 (Yoccoz (1984)). *Soit α 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 α , est conjugué à la rotation R_α par un difféomorphisme h de classe $C^{k-1-\beta-\epsilon}$ pour tout $\epsilon > 0$. \square*

2.4 Exercices

Exercice E.12. 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 λ .

Le but de l'exercice est de donner une classification à conjugaison près des applications strictement dilatantes du cercle dans lui-même.⁵⁸ Nous considérons le cercle comme l'espace métrique $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$, pour la distance d induite⁵⁹ de celle de \mathbb{R} .

(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.

⁵⁸. Nous renvoyons à [Gro] pour la généralisation optimale de ce résultat.

⁵⁹. 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 $d(\hat{x}, \hat{y}) = |x - y|$.

(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 ϕ_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 $\frac{1}{p}$ -contractante⁶⁰ de E dans E .

(5) Conclure.

60. 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 $d(f(x), f(y)) \leq c d(x, y)$.

2.5 Indications pour la résolution des exercices

Correction de l'exercice E.11. Écrivons $\alpha = \frac{p}{q} \bmod \mathbb{Z}$ 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 $\rho : x \mapsto \dot{x}$ 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.

(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 $kp = 1 \bmod q$. Soit $x \in \mathbb{R}$ tel que \dot{x} soit périodique pour f . Les intervalles

$$[x, \tilde{f}(x)[, [\tilde{f}(x), \tilde{f}^2(x)[, \dots, [\tilde{f}^{q-1}(x), \tilde{f}^q(x)[\quad (12)$$

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

$$A = \rho^{-1}(\{\dot{x}, f(\dot{x}), \dots, f^{q-1}(\dot{x})\}).$$

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 (12) 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 $q'p = 1 \bmod q$, 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_α par l'assertion (2). Une conjugaison h entre f et R_α 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.12. (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 λ . 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 ϵ' 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 \tilde{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 ⁶¹ des applications strictement contractantes (nous avons $\frac{1}{p} < 1$), 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 $p h(x) = h \circ f(x) \pmod{\mathbb{Z}}$, donc $\phi_p \circ \bar{h} = \bar{h} \circ f$. D'où \bar{h} est une semi-conjugaison entre (\mathbb{T}^1, ϕ_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 \bar{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 ϕ_p .

61. Ce théorème est le suivant.

Théorème 2.10. (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 ℓ 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 $d(x, y) = d(f(x), f(y)) \leq c d(x, y) < d(x, y)$, 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 $n, p \in \mathbb{N}$,

$$d(x_{n+p}, x_n) \leq \sum_{k=0}^{p-1} d(x_{n+k+1}, x_{n+k}) \leq \frac{c^n}{1-c} d(x_1, x_0). \quad (13)$$

Par conséquent, la suite $(x_n)_{n \in \mathbb{N}}$ est de Cauchy ⁶³ 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 2.10 découle de la formule (13) en prenant $n = 0$ et en faisant tendre p vers $+\infty$.

3 Dynamique linéaire des tores

Fixons $N \in \mathbb{N} - \{0\}$. Le *tore* \mathbb{T}^N est le groupe topologique⁶² 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 θ_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*⁶³ (aussi appelée mesure de Lebesgue) $\lambda = d\theta$, 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 λ découle de l'invariance par translations de la mesure de Lebesgue de \mathbb{R}^N car $\tau_{\dot{y}}(\dot{x}) = \rho(x + y)$ pour tous les x, y dans \mathbb{R}^N . L'unicité de λ 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 $[0, 1]^N$.

Après la partie 2, où nous nous sommes intéressé au cas $N = 1$, nous allons maintenant commencer à étudier des systèmes dynamiques sur le tore \mathbb{T}^N (voir la partie 6.5 pour la suite de l'étude). Cette étude sera facilitée par la possibilité d'utiliser de l'analyse de Fourier⁶³ (voir par exemple [Die2, Chap XIX] pour toute information).

3.1 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

$$\overline{e_m} = e_{-m} \quad \text{et} \quad e_m e_n = e_{m+n} . \tag{14}$$

62. 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 $GL_N(\mathbb{K})$ 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 λ_g et ρ_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 $x \mapsto xg$).



63. Augustin Cauchy
1789–1857



Alfred Haar
1885–1933



Joseph Fourier
1768–1830



Marshall Stone
1903–1989



Karl Weierstrass
1803–1896

Pour toute mesure borélienne μ sur \mathbb{T}^N , notons $(c_m(\mu))_{m \in \mathbb{Z}^N}$ la famille indexée par \mathbb{Z}^N de ses *coefficients de Fourier*, où

$$c_m(\mu) = \int_{\mathbb{T}^N} e_{-m} d\mu = \int_{\mathbb{T}^N} e^{-2i\pi \langle m, \theta \rangle} d\mu(\theta).$$

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 3.1. *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) *L'application $\mu \mapsto (c_m(\mu))_{m \in \mathbb{Z}^N}$ 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*

$$\int_{\mathbb{T}^N} e_m \overline{e_{m'}} d\theta = \int_{\mathbb{T}^N} e_{m-m'} d\theta = \begin{cases} 1 & \text{si } m = m' \\ 0 & \text{sinon.} \end{cases} \quad (15)$$

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 (14)). L'assertion (1) découle donc du théorème de Stone-Weierstrass.⁶⁴

(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. \square

^{64.} Le théorème de Stone⁶³-Weierstrass⁶³ 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.

3.2 Unique ergodicité des translations sur le tore

Le résultat ci-dessous généralise les propriétés déjà vues des rotations du cercle.

Proposition 3.2. 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) τ_α est uniquement ergodique ;
- (2) τ_α est ergodique pour $\lambda = d\theta$;
- (3) τ_α admet une orbite dense ;
- (4) toutes les orbites de τ_α sont denses ;
- (5) τ_α est minimal ;
- (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.1](#).

(3) \Rightarrow (4) Si l'orbite d'un point θ est dense dans \mathbb{T}^N , il en est de même de l'orbite de tous les autres points θ' , car comme les translations τ_α et $\tau_{\theta' - \theta}$ commutent, l'application continue $\tau_{\theta' - \theta}$ envoie l'orbite de θ par τ_α sur celle de θ' .

(4) \Rightarrow (5) Ceci découle de l'équivalence des définitions d'un système dynamique minimal, voir la partie [1.5](#).

(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 θ dans \mathbb{T}^N tels que $\langle m, \theta \rangle = 0$ dans \mathbb{R}/\mathbb{Z} serait donc un fermé non vide invariant par τ_α , différent de \mathbb{T}^N , contradiction.

(6) \Rightarrow (1) Soit μ une mesure de probabilité borélienne invariante par τ_α . 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 décomposition en série de Fourier (voir la proposition [3.1 \(2\)](#)) assure que $\mu = d\theta$. \square

3.3 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 $\mathcal{P} : x \mapsto \dot{x} = x \bmod \mathbb{Z}^N$ la projection canonique

$$\begin{array}{ccc} \mathbb{R}^N & \xrightarrow{M} & \mathbb{R}^N \\ \mathcal{P} \downarrow & & \downarrow \mathcal{P} \\ \mathbb{T}^N & \xrightarrow{\phi_M} & \mathbb{T}^N \end{array} .$$

Lemme 3.3. *Supposons que le déterminant de M soit non nul. Alors*

- (1) *l'application ϕ_M est surjective,*
- (2) *l'application ϕ_M préserve la mesure de Haar $\lambda = d\theta$ sur \mathbb{T}^N ,*
- (3) *pour tout θ dans \mathbb{T}^N , nous avons $\text{Card } \phi_M^{-1}(\theta) = |\det M|$.*

Démonstration. (1) En effet, M et \mathcal{P} sont surjectives.

(2) Comme nous avons $\phi_M \circ \tau_\theta = \tau_{\phi_M(\theta)} \circ \phi_M$ pour tout θ dans \mathbb{T}^N , la mesure de probabilité image $(\phi_M)_*\lambda$ est invariante par $\tau_{\phi_M(\theta)}$. Comme ϕ_M est surjective, et puisque λ est la seule mesure de probabilité invariante par toutes les translations sur \mathbb{T}^N , nous avons $(\phi_M)_*\lambda = \lambda$.

(3) L'application ϕ_M est un morphisme de groupes, donc toutes ses fibres ont le même cardinal. Comme l'application linéaire M est injective et puisque \mathcal{P} 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 $B_\epsilon = \mathcal{P}(A_\epsilon)$ pour tout ϵ dans $]0, (2N \sup_{1 \leq i, j \leq N} |m_{ij}|)^{-1}[$. Nous avons, par la commutativité du diagramme ci-dessus,

$$\mathcal{P}^{-1} \phi_M^{-1} \phi_M \mathcal{P}(A_\epsilon) = M^{-1} \mathcal{P}^{-1} \mathcal{P} M(A_\epsilon) = M^{-1}(MA_\epsilon + \mathbb{Z}^N) = A_\epsilon + M^{-1}\mathbb{Z}^N .$$

Puisque $\phi_M^{-1}(0) = \mathcal{P}(M^{-1}\mathbb{Z}^N)$ et puisque \mathcal{P} 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 ϵ . Par invariance de λ , 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 $(\mathbb{T}^N, \lambda, \phi_M)$. Nous introduisons pour cela une propriété plus forte que l'ergodicité.

Si (X, \mathcal{B}, μ) 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 ⁶⁵ de $\mathbb{L}^2(X, \mu)$ orthogonal aux fonctions 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 3.4. *Soit $(X, \mathcal{B}, \mu, \phi)$ un système dynamique mesuré, avec μ une mesure de probabilité. Les conditions suivantes sont équivalentes.* ⁶⁶

(1) *L'application ϕ est mélangeante, 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) *Nous avons*

$$\forall f, g \in \mathbb{L}^2(X, \mu), \quad \lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g \, d\mu = \left(\int_X f \, d\mu \right) \left(\int_X g \, d\mu \right).$$

(3) *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*

$$\forall f \in E, \forall g \in E', \quad \lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g \, d\mu = \left(\int_X f \, d\mu \right) \left(\int_X g \, d\mu \right).$$

(4) *Nous avons*

$$\forall f, g \in \mathbb{L}_0^2(X, \mu), \quad \lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g \, d\mu = 0.$$

(5) *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 transformation ou le flot ϕ est sous-entendu, nous dirons aussi que la mesure μ est *mélangeante*. Certains auteurs disent *mélange fort* au lieu de mélange, le *mélange faible* signifiant que l'action diagonale de ϕ sur $X \times X$ est ergodique pour $\mu \times \mu$.

(ii) La propriété de mélange signifie que, pour les grands temps t , les événements A et $\phi^{-t}(B)$ sont asymptotiquement indépendants.

(iii) Un système dynamique mesuré mélangeant est ergodique. En effet, lorsque le temps est discret (la démonstration est analogue en temps continu), si $A = \phi^{-1}(A)$, 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.



David Hilbert
65. 1862–1943



Karl Schwarz
1843–1921



Roger Howe
1945–



Calvin Moore
1936–



Elie Cartan
1869–1951

66. Toujours avec la convention que $t \in \mathbb{N}$ lorsque $(X, \mathcal{B}, \mu, \phi)$ est à temps discret

(iv) Notons que par invariance de la mesure, 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 (3) ou de $\mathbb{L}_0^2(X, \mu)$ dans l'assertion (5). Si $\mathbb{L}^2(X, \mu)$ est séparable,⁶⁷ 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 μ 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 $\mathbb{L}^2(X, \mu)$.
- 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 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 3.4. Voici une démonstration par souci de complétude (voir le cours de Théorie ergodique ou [PY, Wal, CorFS, Pet, BenP, EW, Cou]). Rappelons que par l'inégalité de Cauchy-Schwarz⁶⁵, 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.

Notons $P(f, g)$ la propriété $\lim_{t \rightarrow +\infty} \int_X (f \circ \phi^t) g d\mu = \left(\int_X f d\mu \right) \left(\int_X g d\mu \right)$. Elle est vérifiée si f ou si g est constante (car ϕ préserve μ). Par bilinéarité, les assertions (2) et (4), ainsi que les assertions (3) et (5), sont donc équivalentes.

L'assertion (2) implique trivialement l'assertion (3). Elle implique l'assertion (1), car il suffit de prendre les fonctions indicatrices $f = \mathbb{1}_B$ et $g = \mathbb{1}_A$. Si l'assertion (1) est vérifiée, alors par bilinéarité, elle est vérifiée pour f et g des fonctions étagées. Puisque l'espace vectoriel des fonctions étagées est dense dans $\mathbb{L}^2(X, \mu)$, il suffit de démontrer que l'assertion (5) implique l'assertion (2), ou de manière équivalente que l'assertion (5) implique l'assertion (4), pour terminer la démonstration de la proposition 3.4.

⁶⁷. Comme μ est une mesure de probabilité, il suffit pour cela que la σ -algèbre de X soit engendrée par une partie dénombrable (par exemple si X est métrisable séparable), voir par exemple [Coh].

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 $P(f, g)$ est 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 μ , 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 (5) implique l'assertion (4). \square

Exercice E.13. Soient X un espace topologique, $\phi : X \rightarrow X$ une application continue et μ une mesure de probabilité borélienne invariante par ϕ sur X , de support X . Montrer que si le système dynamique mesuré (X, μ, ϕ) est mélangeant, alors le système dynamique topologique (X, ϕ) est topologiquement mélangeant.

Montrer qu'une rotation du cercle n'est pas mélangeante pour la mesure de Lebesgue du cercle.

Voici une propriété plus forte que celle du mélange. Soit $(X, \mathcal{B}, \mu, \phi)$ un système dynamique mesuré, avec μ une mesure de probabilité. 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 ϕ' est une transformation préservant la mesure et commutant avec ϕ , 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, Sar2, 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 ψ -mélangeant sur E , ou qu'il est à ψ -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 ψ -mélangeant sur un sous-espace vectoriel dense E implique être mélangeant, par l'équivalence des définitions dans la proposition 3.4, 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 $t \in \mathbb{R}$.

Exemples. S'il existe $\kappa > 0$ tel que le système dynamique mesuré $(X, \mathcal{B}, \mu, \phi)$ soit ψ -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 ψ -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 ψ -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 mesuré, avec μ une mesure de probabilité et X 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 α -höldériennes sur X , muni de la norme α -höldérienne $f \mapsto \|f\|_{\infty, \alpha}$ (voir la note de bas de page 55). Notons que l'espace vectoriel normé $(C_b^\alpha(X), \|\cdot\|_{\infty, \alpha})$ est un espace de Banach.

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 ℓ 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} .

Soient $p \in [2, +\infty]$ et $k \in \mathbb{N}$. Soit $(X, \mathcal{B}, \mu, \phi)$ un système dynamique mesuré, avec μ une mesure de probabilité et X un ouvert de \mathbb{R}^N ou le tore \mathbb{T}^N muni de leur tribu borélienne. ⁶⁸ 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* ⁶⁹ *de Sobolev*, notée $\|\cdot\|_{p,k}$ ou $\|\cdot\|_{W^{p,k}}$, définie par

$$\|f\|_{p,k} = \|f\|_p + \sum_{\substack{1 \leq \ell \leq k \\ 1 \leq i_1, \dots, i_\ell \leq N}} \|\partial_{i_1, \dots, i_\ell} f\|_p.$$

Théorème 3.5. *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 :*

68. Ceci se généralise à toutes les variétés différentielles.

69. Attention, cette norme sur $E_{p,k}$ n'est pas complète. Le complété de $E_{p,k}$ pour cette norme est noté $W^{p,k}(X)$.

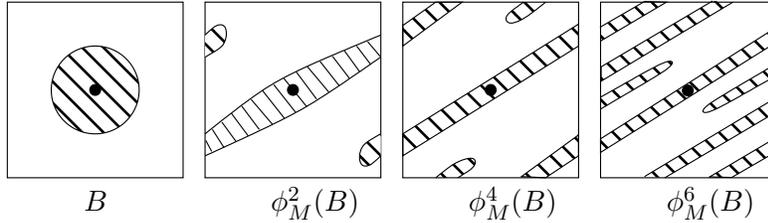
- (1) ϕ_M est mélangeante pour la mesure de Haar $\lambda = d\theta$,
- (2) ϕ_M est ergodique pour la mesure de Haar $\lambda = d\theta$,
- (3) aucune valeur propre (complexe) de M n'est racine de l'unité, ou, de manière équivalente,

$$\forall n \in \mathbb{N} - \{0\}, \quad \det(M^n - \text{id}) \neq 0.$$

De plus, si les valeurs propres (complexes) de M sont de module strictement supérieur à 1 (nous dirons que ϕ_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 3.5 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 ppcm 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 ϕ_M et non constante (car son produit scalaire \mathbb{L}^2 avec e_0 est nul). Donc ϕ_M n'est pas ergodique.

Montrons que l'assertion (3) implique l'assertion (1). Par la proposition 3.4, 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 ℓ, 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

$$e_m \circ \phi_M^n = e_{m_n} . \quad (16)$$

Par hypothèse, tM 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 3.1 (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 $f, g \in C^k(\mathbb{T}^N)$ 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 3.1 (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 (16) et les formules d'orthogonalité (15), 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 $h \in C^{\ell}(\mathbb{T}^N)$, par la proposition 3.1 (4) et par récurrence 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

$$|c_{m'}(h)| \leq \frac{1}{\|m'\|_{\infty}^{\ell}} \|h\|_{\infty, \ell} . \quad (17)$$

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

$$\|m_n\|_{\infty} = \|{}^tM^n m\|_{\infty} \geq \|m\|_{\infty} \inf_{x \in \mathbb{R}^N : \|x\|_{\infty} = 1} \|{}^tM^n x\|_{\infty} \geq c' \lambda^n \|m\|_{\infty} . \quad (18)$$

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 (17) et (18),

$$\begin{aligned} |c_n(f, g)| &\leq \sum_{m \in \mathbb{Z}^N - \{0\}} |c_m(f)| |c_{-m_n}(g)| \leq \sum_{m \in \mathbb{Z}^N - \{0\}} \frac{1}{\|m\|_\infty^k \|m_n\|_\infty^k} \|f\|_{\infty, k} \|g\|_{\infty, k} \\ &\leq \sum_{p \in \mathbb{N} - \{0\}} \sum_{m \in \mathbb{Z}^N - \{0\}, \|m\|_\infty = p} \frac{1}{\|m\|_\infty^k (c' \lambda^n \|m\|_\infty)^k} \|f\|_{\infty, k} \|g\|_{\infty, k} \\ &= \sum_{p \in \mathbb{N} - \{0\}} \frac{2N(2p+1)^{N-1}}{(c')^k p^{2k} \lambda^{kn}} \|f\|_{\infty, k} \|g\|_{\infty, k} = c e^{-\kappa n} \|f\|_{\infty, k} \|g\|_{\infty, k}, \end{aligned}$$

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

3.4 Exercices

Exercice E.14. Soient $(X, \mathcal{B}, \mu, \phi)$ et $(X', \mathcal{B}', \mu', \phi')$ deux systèmes dynamiques mesurés, avec μ et μ' des mesures de probabilité. Soit $h : X \rightarrow X'$ une semi-conjugaison entre (X, ϕ) et (X', ϕ') . Parmi les propriétés suivantes, quelles sont celles qui sont vérifiées par (X', ϕ') si elles le sont par (X, ϕ) , 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.15),
- vérifier la propriété de mélange multiple (voir l'exercice E.15),
- être ψ -mélangeant sur un sous-espace vectoriel dense.

Exercice E.15. Soit $(X, \mathcal{A}, \mu, \phi)$ un système dynamique mesuré, avec μ une mesure de probabilité. Pour k dans $\mathbb{N} - \{0, 1\}$, nous dirons que $(X, \mathcal{A}, \mu, \phi)$ (ou ϕ lorsque (X, \mathcal{A}, μ) est sous-entendu ou μ lorsque (X, \mathcal{A}, ϕ) est sous-entendu) est *mélangeant d'ordre k* ⁷⁰ si pour toutes les parties mesurables A_1, \dots, A_k de X , et toutes les applications τ_1, \dots, τ_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 ϕ lorsque (X, \mathcal{A}, μ) est sous-entendu ou tout simplement μ lorsque (X, \mathcal{A}, ϕ) 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

70. Remarquons que la propriété de mélange d'ordre 2 est équivalente à la propriété de mélange, par invariance de la mesure μ par ϕ . 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 τ_1, \dots, τ_k comme ci-dessus, nous avons

$$\int_X \left(\prod_{i=1}^k f_i \circ \phi^{\tau_i(t)} \right) d\mu \xrightarrow{t \rightarrow +\infty} \prod_{i=1}^k \int_X f_i d\mu. \quad (19)$$

• 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 τ_1, \dots, τ_k comme ci-dessus, la formule (19) soit vérifiée.

(2) Soient \mathcal{A} un alphabet dénombrable (muni de la topologie discrète) et μ 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.16. Soient (X, μ, ϕ) et (X', μ', ϕ') deux systèmes dynamiques mesurés, tous deux à temps continu ou à temps discret, avec μ et μ' des mesures de probabilité. 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, μ, ϕ) et (X', μ', ϕ') 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, μ, ϕ) et (X', μ', ϕ') 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, μ, ϕ) est ψ -mélangeant sur E et si (X', μ', ϕ') est ψ -mélangeant sur E' , alors $(X \times X', \mu \otimes \mu', \phi \times \phi')$ est ψ -mélangeant sur E'' .

Exercice E.17. 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, ϕ_M) est non inversible, que l'ensemble $\text{Per}(\phi_M)$ des points périodiques de ϕ_M est dense dans le tore \mathbb{T}^N , que le système dynamique topologique (\mathbb{T}^N, ϕ_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 $f, g \in C^0(\mathbb{T}^N)$,

$$\int_{\mathbb{T}^N} (f \circ \phi_M) g \, d\lambda = \int_{\mathbb{T}^N} f (L_M g) \, d\lambda. \quad (20)$$

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

$$\left| L_M^n f(x) - \int_{\mathbb{T}^N} f \, d\lambda \right| \leq \frac{c}{\lambda^n} \int_{\mathbb{T}^N} \|d_x f\| \, d\lambda(x). \quad (21)$$

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 ϕ_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 ϕ_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 $\|df\|_\infty = \sup_{x \in \mathbb{T}^N} \|d_x f\|$.

3.5 Indications pour la résolution des exercices

Correction de l'exercice E.13. Soient U et V des ouverts non vides. Puisque la mesure μ est de support total, nous avons $\mu(U), \mu(V) \neq 0$. Si (X, μ, ϕ) 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, ϕ) est topologiquement mélangeant.

Les rotations du cercle ne sont pas topologiquement mélangeantes, comme vu dans l'exemple (1) de la partie 1.5, donc ne sont pas mélangeantes pour la mesure de Lebesgue du cercle.

Correction de l'exercice E.14. • Supposons que $(X, \mathcal{B}, \mu, \phi)$ soit ergodique. Si une partie mesurable $A' \in \mathcal{B}'$ est invariante par ϕ' , 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 ϕ . 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 τ_1, \dots, τ_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 .⁷¹ Soit donc $\kappa > 0$ tel que $\|f \circ h\|_E \leq \kappa \|f\|_{E'}$ pour tout $f \in E'$.

⁷¹. 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 ψ -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 ψ -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 ψ -mélangeant sur E' .

Correction de l'exercice E.15. (1) La démonstration est très proche de celle de la proposition 3.4. Par invariance de la mesure, si une fonction est de classe \mathbb{L}^k , alors sa composée par une puissance quelconque de ϕ l'est encore. Rappelons que le produit de k fonctions de classe \mathbb{L}^k est de classe \mathbb{L}^1 :

$$\left| \int_X f_1 \dots f_k d\mu \right| \leq \prod_{i=1}^k \|f_i\|_{\mathbb{L}^k} = \prod_{i=1}^k \left(\int_X |f_i|^k d\mu \right)^{1/k}. \quad (22)$$

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 μ 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 (22) 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 (σ -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 τ_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.14 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 Δ_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 (19) 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 λ par ϕ_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

$$\sum_{j=1}^k {}^t M^{\tau_j(n)} m_j = 0 \tag{23}$$

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 (23) implique alors

$$\sum_{j=1}^k {}^t T^{\tau_j(n)} v_j = 0 \tag{24}$$

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 (24).

Correction de l'exercice E.16. 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 μ 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,

$$|c_{n,\phi \times \phi'}(f \cdot f', g \cdot g')| = O(\psi(n) \|f\|_E \|g\|_E \|f'\|_{E'} \|g'\|_{E'}).$$

Correction de l'exercice E.17. (1) Par le lemme 3.3 (3), nous savons que les fibres $\phi_M^{-1}(\theta)$ pour θ dans \mathbb{T}^N sont de cardinal égal à $|\det M| > 1$, donc ϕ_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 4 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 3.5 que la transformation ϕ_M du tore est mélangeante pour la mesure de Haar de \mathbb{T}^N , donc est topologiquement mélangeante par l'exercice E.13, donc est positivement transitive et non errante comme vu dans la partie 1.5.

(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 3.3 (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^∞ -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

$$\begin{aligned} \int_{\mathbb{T}^N} (f \circ \phi_M) g \, d\lambda &= \sum_{[i] \in \mathbb{Z}^N / M\mathbb{Z}^N} \int_{M^{-1}([0, 1]^N + i)} \tilde{f}(Mx) \tilde{g}(x) \, dx \\ &= \sum_{[i] \in \mathbb{Z}^N / M\mathbb{Z}^N} \int_{[0, 1]^N + i} \tilde{f}(y + i) \tilde{g}(M^{-1}(y + i)) \frac{1}{|\det M|} \, dy \\ &= \int_{\mathbb{T}^N} f(L_M g) \, d\lambda. \end{aligned}$$

4 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 précédente 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 $G = \mathrm{SL}_N(\mathbb{R})$.⁷²

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 4.5, 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 , seront remplacées ici par les représentations unitaires de G , voir la partie 4.2.

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 $(t, f) \mapsto (\pi(t)f : x \bmod \mathbb{T}^N \mapsto f(x + t \bmod \mathbb{T}^N))$. 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 3.2) 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 4.9).

4.1 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 [Paul1, 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

$$d_G(gx, gy) = d_G(x, y).$$

Par exemple, en notant indifféremment $e = \mathrm{id} = I_N$ la matrice identité de taille N et $\| \cdot \|$ une norme matricielle⁷³ quelconque sur $\mathcal{M}_N(\mathbb{R})$, nous pouvons prendre la distance⁷⁴

$$d_G(x, y) = \ln(1 + \|x^{-1}y - \mathrm{id}\|) + \ln(1 + \|y^{-1}x - \mathrm{id}\|). \quad (25)$$

72. Cette étude se généralise à tous les groupes de Lie réels G connexes, semi-simples, de centre fini, voir par exemple [BekM].

73. c'est-à-dire une norme sur $\mathcal{M}_N(\mathbb{R})$ vérifiant $\|xy\| \leq \|x\| \|y\|$, comme la norme d'opérateur des matrices vues comme applications linéaires de l'espace euclidien standard \mathbb{R}^N

74. L'inégalité triangulaire découle par exemple des inégalités

$$\begin{aligned} 1 + \|x^{-1}z - e\| &= 1 + \|x^{-1}y - e\| + \|y^{-1}z - e\| + (\|x^{-1}z - e\| - \|x^{-1}y - e\| - \|y^{-1}z - e\|) \\ &\leq 1 + \|x^{-1}y - e\| + \|y^{-1}z - e\| + \|(x^{-1}z - e) - (x^{-1}y - e) - (y^{-1}z - e)\| \\ &= 1 + \|x^{-1}y - e\| + \|y^{-1}z - e\| + \|(x^{-1}y - e)(y^{-1}z - e)\| \leq (1 + \|x^{-1}y - e\|)(1 + \|y^{-1}z - e\|). \end{aligned}$$

Lemme 4.1. (1) *Il existe une mesure borélienne positive non nulle ν_G sur G , finie sur les compacts, invariante par les translations à gauche⁷⁵, unique à un scalaire multiplicatif près.*

(2) *La mesure μ est aussi invariante par les translations à droite⁷⁶.*

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 ν_G est appelée une *mesure de Haar* (à gauche) de G . Lorsque l’assertion (2) est vérifiée, nous disons 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 ν_G pour $G = \mathrm{SL}_N(\mathbb{R})$. Notons λ la mesure de Lebesgue sur $\mathcal{M}_N(\mathbb{R}) = \mathbb{R}^{N^2}$, et posons, pour tout borélien A de G ,

$$\nu_{\mathrm{SL}_N(\mathbb{R})}(A) = \lambda\left(\left\{x \in \mathcal{M}_N(\mathbb{R}) : 1 \leq \det x \leq 2 \quad \text{et} \quad \frac{1}{(\det x)^{\frac{1}{N}}} x \in A\right\}\right).$$

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 .⁷⁷

Pour montrer l’unicité, soient ν et ν' 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 $G = \bigcup_{g \in P} g B_G(e, \epsilon)$. Puisque ν 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 σ -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.

Fixons $g \in C_c^0(G)$ positive non identiquement nulle⁷⁸. 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).

75. 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

76. 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

77. 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 $(-1)^{\frac{n(n-1)}{2}}$.

78. par exemple la restriction à G d’une fonction dans $C_c^0(\mathbb{R}^{N^2})$, positive et valant 1 en I_N

Par le théorème de Fubini (puisque les mesures sont σ -finies) et par l'invariance par translation à gauche de ν par y^{-1} et de ν' 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 ν , 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 ν_G sur G . Soit Γ 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 ρ la projection canonique de G dans $X = \Gamma \backslash G$ définie par $g \mapsto \Gamma g$.

Nous notons encore ρ l'action de G sur X par translations à droite sur les classes à gauche modulo Γ : 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 4.2. (1) Pour tout g dans G , l'application ρ 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 γ dans Γ .

(2) Il existe une mesure positive borélienne non nulle $\mu = \mu_X$ sur X , finie sur les compacts et invariante par translations à droite⁷⁹, unique à scalaire multiplicatif près.

79. c'est-à-dire telle que $(\rho_g)_* \mu = \mu$ pour tout $g \in G$

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 Γ . La mesure μ 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 Γ est remplacé par un sous-groupe fermé de G). Attention, la mesure μ n’est en général pas la mesure image de ν_G par \mathcal{r} . Par contre, la restriction de μ à toute boule $B(\Gamma g, \frac{\epsilon_0}{2})$ est la mesure image par \mathcal{r} de la restriction d’un multiple scalaire strictement positif de ν_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 \mathcal{r} est un revêtement (voir par exemple [Pau4]).

(2) Nous définissons d’abord μ 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 $A = \mathcal{r}(\tilde{A})$ 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 ν_G est invariante à gauche.

Définissons alors μ 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 Γ de G est appelé un *réseau* si $\mu(\Gamma \backslash G) < \infty$, ce qui ne dépend pas du choix de μ . Nous normaliserons alors la mesure de Haar μ de l’espace homogène $\Gamma \backslash G$ pour qu’elle soit de probabilité. Un sous-groupe discret Γ 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 4.16 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].

4.2 Représentations unitaires

Soit Γ 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 μ 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),⁸⁰ notons $U(\mathcal{H})$ le *groupe unitaire* de \mathcal{H} , c’est-à-dire le groupe des bijections linéaires continues⁸¹ $u : \mathcal{H} \rightarrow \mathcal{H}$

- qui préservent le produit scalaire de \mathcal{H} :

$$\forall x, y \in \mathcal{H}, \quad \langle u(x), u(y) \rangle = \langle x, y \rangle,$$

80. Son produit scalaire sera noté $(x, y) \mapsto \langle x, y \rangle$. Il est supposé linéaire à gauche et anti-linéaire à droite.

81. leur inverse est alors automatiquement continu par le théorème de Banach, mais cela découle aussi de la propriété suivante d’isométrie

- 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'adjoint⁸² de u .

Une *représentation unitaire* de G dans un espace de Hilbert \mathcal{H} est un morphisme de groupes $\pi : G \rightarrow \mathrm{U}(\mathcal{H})$ 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 $\pi : G \rightarrow \mathrm{U}(\mathbb{L}^2(X, \mu))$ 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

$$\mathbb{L}_0^2(X, \mu)^G = \{0\} . \tag{26}$$

Démonstration. Montrons que π est en effet une représentation unitaire. Comme μ 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 $\mathrm{U}(\mathcal{H})$. Il est immédiat que π est un morphisme de groupes.⁸³ 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

$$\|\pi(g_n)f - \pi(g)f\|_{\mathbb{L}^2} \leq 2\|f - f_0\|_{\mathbb{L}^2} + \|\pi(g_n)f_0 - \pi(g)f_0\|_{\mathbb{L}^2} \leq 3\epsilon$$

82. 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 $x, y \in \mathcal{H}$.

83. 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 $(\pi(gg')f)(x) = f(xgg') = (\pi(g')f)(xg) = ((\pi(g)\pi(g')f)(x))$.

pour n assez grand.

Comme la mesure μ 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 μ -presque tout x . Donc, par le théorème de Fubini, pour μ -presque tout x , nous avons $f(xg) = f(x)$ pour ν_G -presque tout g . Fixons un tel point x . Lorsque g décrit le complémentaire d'un ensemble ν_G -négligeable, les points xg décrivent le complémentaire d'un ensemble μ -négligeable. Donc f est constante μ -presque tout. Comme f est d'intégrale nulle, elle est nulle μ -presque tout. \square

Le résultat suivant découle de ce qui précède.

Corollaire 4.3. *Si μ est une mesure de Haar sur l'espace homogène $X = \Gamma \backslash G$, alors $\mathbb{L}^2(X, \mu)^G$ est réduit aux fonctions μ -presque partout constantes.* \square

Soit $\pi : G \rightarrow \mathbb{U}(\mathcal{H})$ 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 π . 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' :

$$\forall g, h, h' \in G, \quad \text{si } \pi(h)v = v \text{ et } \pi(h')v' = v', \quad \text{alors } c_{v,v'}(h'gh) = c_{v,v'}(g). \quad (27)$$

4.3 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 Γ 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 4.4. *Soient Γ 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 μ de l'espace homogène X .*

Démonstration. D'après l'assertion (4) de la proposition 1.6, ce résultat est une conséquence de la proposition suivante, appliquée à $\mathcal{H} = \mathbb{L}^2(X, \mu)$ et π la représentation quasi-régulière de G sur \mathcal{H} , en utilisant le corollaire 4.3. \square

Proposition 4.5. Soient (\mathcal{H}, π) 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 4.6. (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

$$\lim_{n \rightarrow \infty} g_n = g \quad \text{et} \quad \lim_{n \rightarrow \infty} s_n g_n s'_n = \text{id} .$$

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

$$\|\pi(g)v - v\|^2 = 2(\|v\|^2 - \text{Re}(c_{v,v}(g))) .$$

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 $\text{Re}(c_{v,v}(g)) = \|v\|^2$. Or par l'inégalité de Cauchy-Schartz, nous avons

$$\sqrt{\text{Re}(c_{v,v}(g))^2 + \text{Im}(c_{v,v}(g))^2} = |c_{v,v}(g)| \leq \|\pi(g)v\| \|v\| = \|v\|^2 .$$

Donc $\text{Im}(c_{v,v}(g)) = 0$ 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 (27). Par continuité, nous en déduisons que $c_{v,v}(g) = c_{v,v}(\text{id}) = \|v\|^2$. Il résulte de l'assertion b) que g appartient à G_v . \square

Le phénomène de Mautner permet de montrer le lemme suivant.

Lemme 4.7. 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 4.6 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 4.6 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 4.6 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 4.5. 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 4.7 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 engendrent⁸⁴ G , nous avons $G_v = G$. \square

4.4 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.

Théorème 4.8. Soient Γ un réseau de $G = \text{SL}_N(\mathbb{R})$ 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 μ 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, μ, ϕ) est mélangeant si et seulement si son inverse (Y, μ, ϕ^{-1}) l'est, le théorème 4.8 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 Howe⁶⁵-Moore⁶⁵ de décroissance des coefficients des représentations unitaires de G . Le théorème de Howe-Moore, ainsi que le théorème 4.8, sont valables lorsque G est n'importe quel groupe de Lie réel connexe, semi-simple, de centre fini. Voir par exemple [BekM, Chap. III]).

84. 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 précédant 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 4.14.

Proposition 4.9. (Théorème de Howe-Moore) Soient $G = \mathrm{SL}_N(\mathbb{R})$, π 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 4.8. Appliquons le théorème de Howe-Moore 4.9 avec $\mathcal{H} = \mathbb{L}_0^2(\Gamma \backslash G, \mu)$ et π la restriction à \mathcal{H} de la représentation quasi-régulière de G . L'hypothèse de la proposition 4.9 est vérifiée par la formule (26). Le théorème 4.8 découle alors de la proposition 3.4 (4) car pour tous les $v, w \in \mathcal{H}$, le n -ème coefficient de corrélation $c_n(v, w)$ est un coefficient matriciel de π :

$$c_n(v, w) = \int_X (v \circ (\rho_g)^n) w d\mu = \int_X v(xg^{-n}) w(x) d\mu(x) = \langle \pi(g^{-n})v, \bar{w} \rangle. \quad \square$$

Pour montrer le théorème de Howe-Moore 4.9, nous aurons besoin de deux résultats classiques. Notons A le sous-groupe de Cartan⁶⁵ 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 $\mathrm{SL}_N(\mathbb{R})$.

Lemme 4.10. (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. Rappelons⁸⁵ qu'une application linéaire continue $u : \mathcal{H} \rightarrow \mathcal{H}$ est faiblement continue : si $v_n \rightharpoonup v$, alors $u(v_n) \rightharpoonup u(v)$.

85. En effet, pour tout $w \in \mathcal{H}$, nous avons

$$\lim_{n \rightarrow \infty} \langle u(v_n), w \rangle = \lim_{n \rightarrow \infty} \langle v_n, u^*(w) \rangle = \langle v, u^*(w) \rangle = \langle u(v), w \rangle.$$

Lemme 4.11. (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.*⁸⁶

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 Riesz⁸⁷, 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 4.9. 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

$$\forall n \in \mathbb{N}, \quad |\langle \pi(g_n)v', w' \rangle| \geq \epsilon \quad \text{et} \quad \lim_{n \rightarrow \infty} \|g_n\| = \infty .$$

Par le lemme 4.11, 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 4.12. *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 π 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

$$T_1 = |\langle \pi(ka_n k')v' - \pi(ka_n k'_n)v', w'' \rangle| \leq \|\pi(k')v' - \pi(k'_n)v'\| \|w''\| \xrightarrow[n \rightarrow \infty]{} 0 ,$$

86. 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].

87. 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 $(\lambda, x) \mapsto \bar{\lambda} x$). 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 [Rud] ou [Pau3, §6.2].

$$\begin{aligned}
T_2 &= |\langle \pi(k a_n k'_n) v' - \pi(k_n a_n k'_n) v', w'' \rangle| = |\langle \pi(a_n k'_n) v', \pi(k^{-1}) w'' - \pi(k_n^{-1}) w'' \rangle| \\
&\leq \|v'\| \|\pi(k^{-1}) w'' - \pi(k_n^{-1}) w''\| \xrightarrow{n \rightarrow \infty} 0,
\end{aligned}$$

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 :

$$\begin{aligned}
U_{ij} &= \{\text{id} + s E_{ij} : s \in \mathbb{R}\} \\
A_{ij} &= \{\text{id} + (t-1) E_{ii} + (t^{-1}-1) E_{jj} : t > 0\} \\
S_{ij} &= \{\text{id} + (a-1) E_{ii} + b E_{ij} + c E_{ji} + (d-1) E_{jj} : ad - bc = 1\}
\end{aligned}$$

Le sous-groupe U_{ij} est abélien unipotent (isomorphe au groupe additif $(\mathbb{R}, +)$ par l'isomorphisme $s \mapsto \text{id} + s E_{ij}$). 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 $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \text{id} + (a-1) E_{ii} + b E_{ij} + c E_{ji} + (d-1) E_{jj}$). Nous pourrions lui appliquer la proposition 4.5.

Lemme 4.13. *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 $X^{-1} Y X$ 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 ℓ assure que nous avons $\lim_{n \rightarrow \infty} a_n^{-1} u a_n = e$. Par un calcul analogue à celui du lemme 4.12, 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)

$$\pi(u) v_0 = \lim_{n \rightarrow \infty} \pi(u) \pi(a_n) v = \lim_{n \rightarrow \infty} \pi(a_n) \pi(a_n^{-1} u a_n) v = \lim_{n \rightarrow \infty} \pi(a_n) v = v_0.$$

C'est ce que nous voulions. □

Lemme 4.14. *Le groupe $G = \text{SL}_N(\mathbb{R})$ 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 $(\text{id} + s E_{ij}) g$ 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 $g(\text{id} + s E_{ij})$ 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 84. \square

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 4.5, 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 4.5, ce vecteur est invariant par $U_{i'j'}$. Le lemme 4.14 ci-dessus montre alors que le vecteur non nul v_0 est invariant par G , contradiction. \square

4.5 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 quotient⁸⁸ $G = \mathrm{PSL}_2(\mathbb{R})$ 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 $ad - bc = 1$.

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.⁸⁹

Pour tout sous-groupe discret Γ 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 $\mathcal{r} : \mathrm{SL}_2(\mathbb{R}) \rightarrow \mathrm{PSL}_2(\mathbb{R})$ est la projection canonique, alors $\tilde{\Gamma} = \mathcal{r}^{-1}(\Gamma)$ est un sous-groupe discret de $\mathrm{SL}_2(\mathbb{R})$, et \mathcal{r} 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 Γ 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 Γ 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*

$$\mathbb{H} = \{z = x + iy \in \mathbb{C} : y = \mathrm{Im} z > 0\}$$

⁸⁸. Notons que la notation G dans cette partie ne désigne pas le même objet que les notations G des parties 4.1 à 4.4.

⁸⁹. 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 $\mathrm{SL}_2(\mathbb{R})$.

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 l'on a bien $g \cdot (g' \cdot z) = (gg') \cdot z$; et le fait que G préserve \mathbb{H} découle du calcul

$$\operatorname{Im}(g \cdot z) = \frac{\operatorname{Im} z}{|cz + d|^2}. \quad (28)$$

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

$$g \cdot (z, \xi) = \left(\frac{az + b}{cz + d}, \frac{\xi}{(cz + d)^2} \right). \quad (29)$$

Pour tout vecteur tangent $v = (z, \xi) \in T\mathbb{H}$ où $z = x + iy$, notons

$$\|v\|_z = \frac{|\xi|}{y},$$

et appelons cette quantité la *norme hyperbolique* de v .⁹⁰ 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 de vérifier que l'action de G préserve la norme hyperbolique⁹¹, c'est-à-dire que pour tous les $g \in G$ et $v = (z, \xi) \in T\mathbb{H}$, nous avons

$$\|g \cdot v\|_{g \cdot z} = \|v\|_z.$$

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 montre que l'action de G sur $\mathbb{H}_{\mathbb{R}}^2$ préserve le volume hyperbolique.

Notons

$$T^1\mathbb{H}_{\mathbb{R}}^2 = \{v = (z, \xi) \in \mathbb{H}_{\mathbb{R}}^2 \times \mathbb{C} : \|v\|_z = 1\},$$

90. 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].

91. On dit que l'action de G est *isométrique* pour la métrique riemannienne susdite.

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

$$T_z^1\mathbb{H}_{\mathbb{R}}^2 = \{v = (z, \xi) : |\xi| = \text{Im } z\} .$$

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 $\theta \mapsto (z, (\text{Im } z) e^{2i\pi\theta})$, notée $\text{vol}_{T_z^1\mathbb{H}_{\mathbb{R}}^2}(v)$. 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*⁹² sur le fibré tangent unitaire $T^1\mathbb{H}_{\mathbb{R}}^2$, notée m_{Liou} est la mesure qui désintègre par la projection p au-dessus du volume hyperbolique, avec mesures conditionnelles sur les fibres les mesures sphériques :

$$m_{\text{Liou}}(v) = \int_{z \in \mathbb{H}_{\mathbb{R}}^2} \text{vol}_{T_z^1\mathbb{H}_{\mathbb{R}}^2}(v) d\text{vol}_{\mathbb{H}_{\mathbb{R}}^2}(z) .$$

Ceci signifie que pour toute fonction continue à support compact $f \in C_c^0(T^1\mathbb{H}_{\mathbb{R}}^2)$, nous avons

$$\begin{aligned} \int_{T^1\mathbb{H}_{\mathbb{R}}^2} f(v) dm_{\text{Liou}}(v) &= \int_{z \in \mathbb{H}_{\mathbb{R}}^2} \left(\int_{v \in T_z^1\mathbb{H}_{\mathbb{R}}^2} f(v) d\text{vol}_{T_z^1\mathbb{H}_{\mathbb{R}}^2}(v) \right) d\text{vol}_{\mathbb{H}_{\mathbb{R}}^2}(z) \\ &= \int_{x \in \mathbb{R}} \int_{y \in]0, +\infty[} \int_{\theta \in \mathbb{R}/\mathbb{Z}} f(x + iy, y e^{2i\pi\theta}) \frac{d\theta dx dy}{y^2} . \end{aligned} \quad (30)$$

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 4.15. 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 Φ est un homéomorphisme.⁹³
- (2) L'application Φ 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 Φ^{-1} est une mesure de Haar de G .



Joseph Liouville
92. 1809–1882



Shrikrishna Dani

93. en fait un C^∞ -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

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

$$\mathrm{PSO}_2(\mathbb{R}) = \mathrm{SO}_2(\mathbb{R})/\{\pm \mathrm{id}\} = \left\{ \rho_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} : \theta \in \mathbb{R}/2\pi\mathbb{Z} \right\}.$$

Comme $T_i^1\mathbb{H}_{\mathbb{R}}^2 = \{(i, \xi) : |\xi| = 1\}$ et puisque par la formule (29) 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 $\mathrm{PSO}_2(\mathbb{R})$ agit simplement transitivement sur $T_i^1\mathbb{H}_{\mathbb{R}}^2$. Donc G agit simplement transitivement sur le fibré tangent unitaire $T^1\mathbb{H}_{\mathbb{R}}^2$.

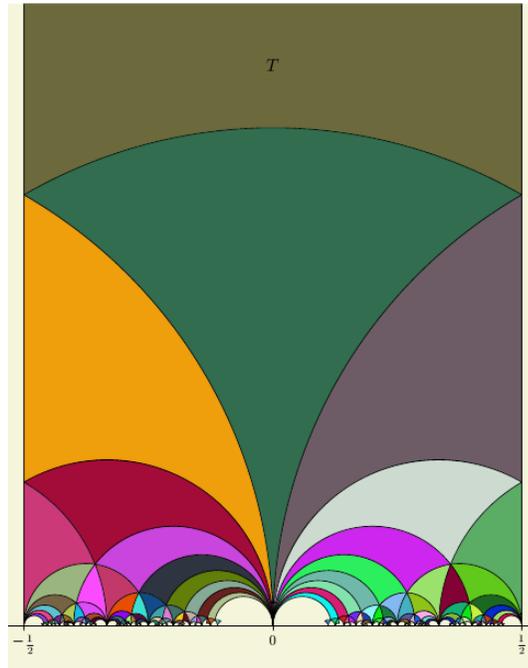
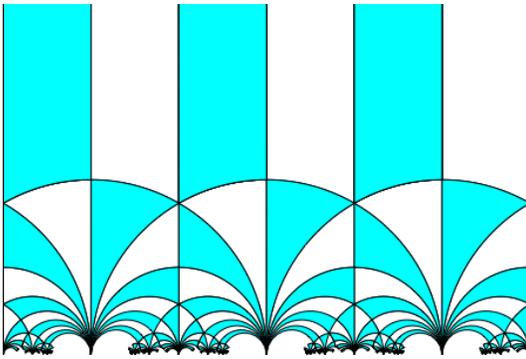
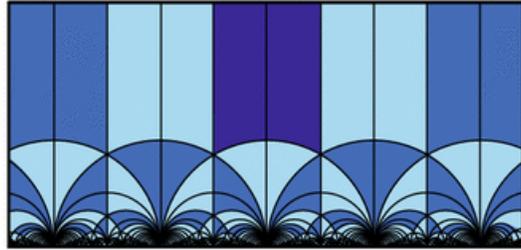
Par conséquent, l'application orbitale $\Phi : g \mapsto g \cdot (i, i)$ est une bijection continue équivariante. Il n'est pas difficile de montrer qu'elle est propre, donc est un homéomorphisme.

L'assertion (3) découle alors des assertions (1) et (2), car celles-ci montrent que la mesure image $(\Phi^{-1})_* m_{\mathrm{Liou}}$ 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

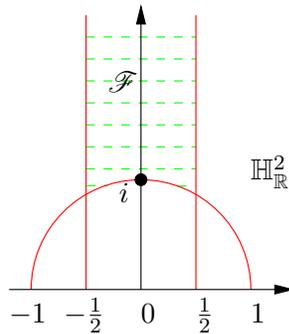
Soit Γ un sous-groupe discret de G . L'homéomorphisme équivariant Φ induit par passage au quotient un homéomorphisme de $\Gamma \backslash G$ dans $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$, que nous notons encore Φ . La mesure de Liouville sur $T^1\mathbb{H}_{\mathbb{R}}^2$ induit une mesure sur $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$, 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 4.2 et la remarque qui le suit, ainsi que le début de la partie 4.5).

Proposition 4.16. *Le sous-groupe discret $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, appelé le groupe modulaire, est un réseau de $G = \mathrm{PSL}_2(\mathbb{R})$.*

Ceci montre aussi que $\mathrm{SL}_2(\mathbb{Z})$ est un réseau de $\mathrm{SL}_2(\mathbb{R})$. 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 4.15, 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

$$\mathcal{F} = \left\{ z \in \mathbb{H} : |z| \geq 1, \quad |\operatorname{Re} z| \leq \frac{1}{2} \right\}$$

est un *domaine (faiblement) fondamental*⁹⁴ pour l'action de Γ sur $\mathbb{H}_{\mathbb{R}}^2$. Voir l'exercice E.19.

Par désintégration de la mesure de Liouville sur $T^1\mathbb{H}_{\mathbb{R}}^2$, et puisque les mesures conditionnelles sur les fibres sont des mesures de probabilité, nous avons⁹⁵

$$m_{\text{Liou}}(\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2) \leq m_{\text{Liou}}(T^1\mathcal{F}) \leq \operatorname{vol}_{\mathbb{H}_{\mathbb{R}}^2}(\mathcal{F}).$$

Or⁹⁶

$$\operatorname{vol}_{\mathbb{H}_{\mathbb{R}}^2}(\mathcal{F}) = \int_{\mathcal{F}} \frac{dx dy}{y^2} \leq \int_{y=\frac{1}{2}}^{\infty} \int_{x=-\frac{1}{2}}^{\frac{1}{2}} \frac{dx dy}{y^2} \leq 2.$$

Ceci montre le résultat. \square

Nous définissons la *longueur (hyperbolique)* $\ell(c)$ d'une courbe $c : [a, b] \rightarrow \mathbb{H}_{\mathbb{R}}^2$, de classe C^1 par morceaux, par la formule

$$\ell(c) = \int_a^b \|c'(t)\|_{c(t)} dt.$$

Une courbe $c : I \rightarrow \mathbb{H}_{\mathbb{R}}^2$ de classe C^1 , où I est un intervalle 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 est facile de vérifier que d est bien une distance, et 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.

94. c'est-à-dire que $\bigcup_{\gamma \in \Gamma} \gamma \mathcal{F} = \mathbb{H}_{\mathbb{R}}^2$. Nous avons aussi, en notant $\overset{\circ}{\mathcal{F}} = \{z \in \mathbb{H} : |z| > 1, \quad |\operatorname{Re} z| < \frac{1}{2}\}$ 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.19), mais nous n'en aurons pas besoin ici. Il est aussi possible de montrer que

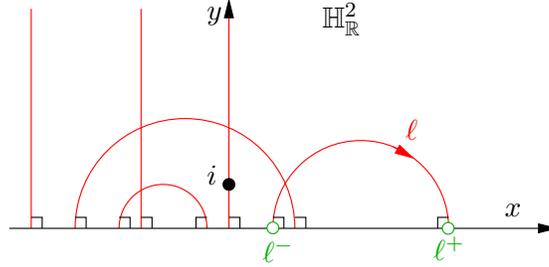
$$\mathcal{F}' = \left\{ z \in \mathcal{F} : \operatorname{Re} z \neq \frac{1}{2} \text{ et } z \neq e^{i\theta} \text{ pour tout } \theta \in \left[\frac{\pi}{3}, \frac{\pi}{2} \right] \right\}$$

est un *domaine fondamental strict*, c'est-à-dire contenant un et un seul point de toute orbite de Γ dans $\mathbb{H}_{\mathbb{R}}^2$.

95. 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 Γ 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.19), du fait déjà mentionné que les images de l'intérieur du domaine fondamental \mathcal{F} par les éléments du groupe dénombrable Γ sont deux à deux disjointes, et parce que la frontière de \mathcal{F} est de mesure nulle.

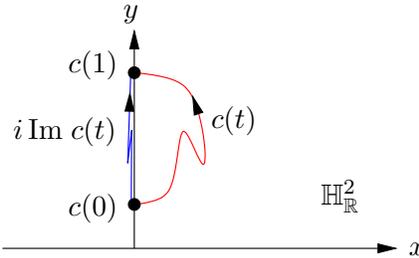
96. Un calcul plus précis laissé au lecteur montre que $\operatorname{vol}_{\mathbb{H}_{\mathbb{R}}^2}(\mathcal{F}) = \frac{\pi}{3}$.

Lemme 4.17. *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^∞ 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 ∞ 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'infinis des rayons géodésiques 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 $|c'(t)| \geq |(\text{Im } c)'(t)|$, nous avons $\ell(c) \geq \ell(i \text{Im } c)$, avec inégalité stricte si la courbe c n'est pas contenue dans l'axe imaginaire.



De plus, tout chemin sur l'axe imaginaire faisant un aller-retour non trivial est de longueur (hyperbolique) strictement supérieure à celle de ce chemin privé de cet aller-retour. Donc, pour tout $y > 1$,

$$d(i, yi) = \int_0^y \frac{ds}{s} = \ln y ,$$

et la courbe $c : [0, \ln y] \rightarrow \mathbb{H}_{\mathbb{R}}^2$ définie par

$$t \mapsto e^t i \tag{31}$$

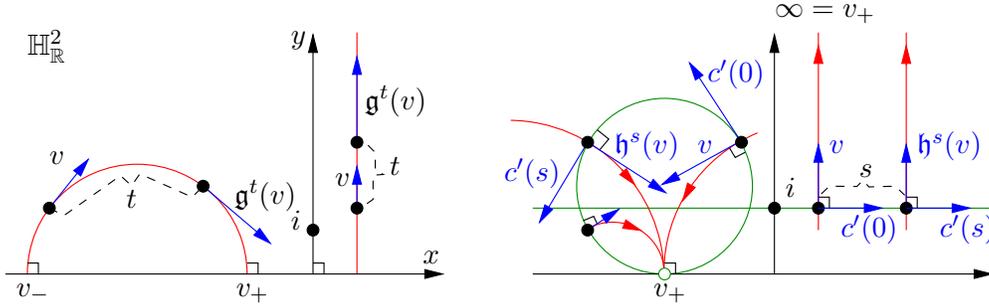
est l'unique (modulo translation à la source) géodésique C^1 par morceaux joignant i à yi . Par approximation, nous montrons facilement que c'est aussi la seule géodésique (continue) joignant i à yi (et elle est donc de classe C^∞). Comme tout élément g de G est une isométrie hyperbolique, la courbe $t \mapsto g \cdot (e^t i)$ est l'unique (modulo translation à la source) géodésique joignant $g \cdot i$ à $g \cdot (yi)$.

Il est bien connu (voir [Rud]) 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 $T^1\mathbb{H}_{\mathbb{R}}^2$, 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 $t \in \mathbb{R}$,

$$\mathfrak{g}^t(v) = (c(t), c'(t)) ,$$

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.



Appelons *horocycles*⁹⁷ les droites horizontales contenues dans $\mathbb{H}_{\mathbb{R}}^2$, dites *centrées* en ∞ , 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 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 ∞ , et les homothéties $z \mapsto \lambda^2 z$ avec $\lambda > 0$ agissent transitivement sur les horocycles centrés en ∞ . 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 par transitivité pour les autres. L'application

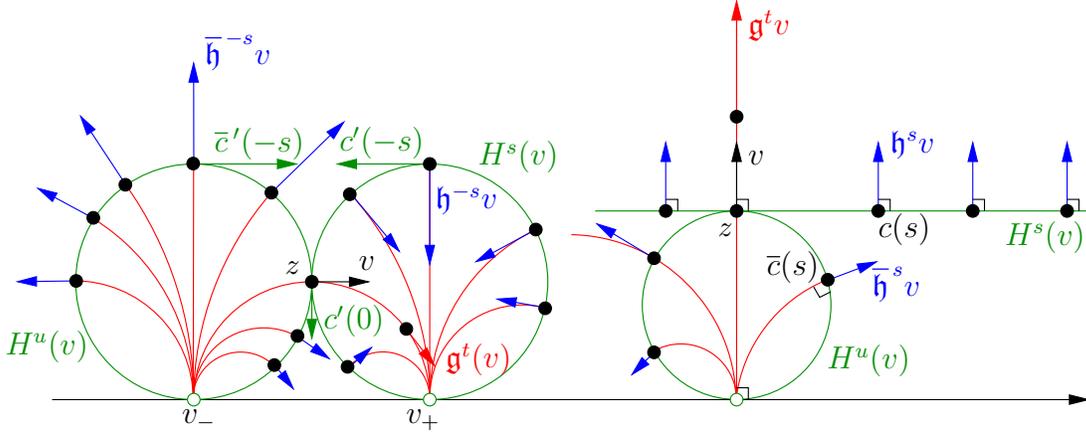
$$c : s \mapsto i + s \tag{32}$$

est l'unique paramétrage 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 *horocyclique 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 *horocyclique 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

$$g \cdot H^s(v) = H^s(g \cdot v) \quad \text{et} \quad g \cdot H^u(v) = H^u(g \cdot v) .$$

97. 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)

$$\mathfrak{h}^s(v) = (c(s), i c'(s)) \in T^1\mathbb{H}_{\mathbb{R}}^2, \quad (33)$$

où $s \mapsto c(s)$ est l'unique paramétrage 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

$$\bar{\mathfrak{h}}^s(v) = (\bar{c}(s), i \bar{c}'(s)) \in T^1\mathbb{H}_{\mathbb{R}}^2, \quad (34)$$

où $s \mapsto \bar{c}(s)$ est l'unique paramétrage par longueur d'arc hyperbolique de l'horocycle instable $H^u(v)$ de v tel que $v = (\bar{c}(0), i \bar{c}'(0))$.

Lemme 4.18. *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^∞ -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 commutent⁹⁸ 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*

$$\mathfrak{g}^t(\Phi(g)) = \Phi(g a_t) \quad \text{et} \quad \mathfrak{h}^s(\Phi(g)) = \Phi(g u_s) \quad (35)$$

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 (31), donc les deux termes de cette égalité valent $(e^t i, e^t i)$, et que

- l'égalité $\mathfrak{h}^s(i, i) = \Phi(u_s)$, ce qui est clair car l'unique paramétrage 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 ce

98. 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 $\mathfrak{h}^s(g \cdot v) = g \cdot \mathfrak{h}^s(v)$

paramétrage, est $s \mapsto i + s$ par la formule (32), donc les deux termes de cette égalité valent $(i + s, i)$.

La dernière assertion découle de la deuxième, du lemme 4.15 (3), et de l'invariance par translations à droite des mesures de Haar de G . \square

Pour tout sous-groupe discret Γ 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 4.18, 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 Γ 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 4.4, est C^∞ -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 4.19. *Si Γ 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 Γ 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 Θ 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 4.15, 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 4.8. \square

Le résultat suivant découle du fait que mélangeant implique ergodique (voir la remarque (iii) suivant la proposition 3.4), 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.1.

Corollaire 4.20. *Si Γ 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_{Liou} de $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$. L'orbite de m_{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.18. *Soit Γ 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*

$$\mathfrak{g}^t \circ \mathfrak{h}^s \circ \mathfrak{g}^{-t} = \mathfrak{h}^{se^{-t}} . \quad (36)$$

(2) Montrer que la famille $(\bar{h}^s)_{s \in \mathbb{R}}$ définie par la formule (34) est un groupe à un paramètre de C^∞ -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 $s, t \in \mathbb{R}$,

$$g^t \circ \bar{h}^s \circ g^{-t} = \bar{h}^{se^t} \quad \text{et} \quad \bar{h}^s(\Phi(g)) = \Phi(gu_s^-)$$

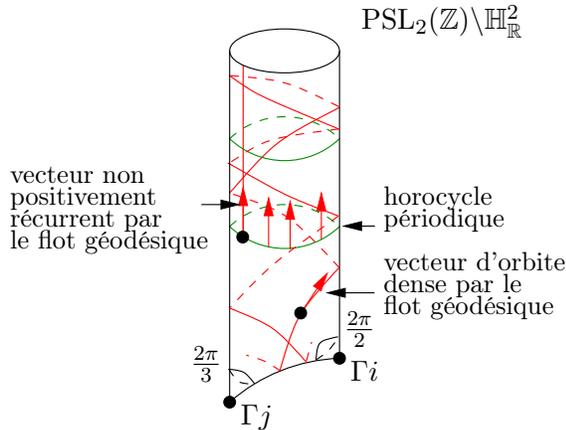
où $u_s^- = \begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix}$. Montrer que si Γ est un réseau de $\mathrm{PSL}_2(\mathbb{R})$, alors le flot horocyclique instable $(\bar{h}^s)_{s \in \mathbb{R}}$ sur $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ défini par passage au quotient, est mélangeant pour la mesure de Liouville de $\Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$.

(3) Si Γ 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 (36), et en montrant que f est invariante par l'action par translations à droite de G sur Y identifié à $\Gamma \backslash G$ par Φ^{-1} .

4.6 Exercices

Exercice E.19. Notons $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ le groupe modulaire, $M = \Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$ la courbe modulaire⁹⁹ et $T^1M = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$.¹⁰⁰ 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 Γ 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 Γ , 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

99. 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.

100. 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 Γ 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 Γi 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 géodésiques périodiques de M et 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 α^σ est le conjugué de Galois de α (l'autre racine d'un polynôme quadratique à coefficients entiers dont α 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 α .
- (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 positivement récurrente 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_{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 $(\Gamma \tilde{v}) \cap (\mathcal{F} \times \mathbb{S}_1) = \{(z, \xi)\}$.¹⁰¹ Montrer, en utilisant le théorème ergodique de Birkhoff 1.7, que pour m_{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.20. 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 Γ est le groupe modulaire, le *théorème de Dani⁹²-Furstenberg* suivant : si Γ 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 sont :

- la mesure de Liouville m_{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*

^{101.} 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 $z \in \mathcal{F}$.

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 $s \mapsto \mathbf{u}^s(v)$, renormalisée pour être de probabilité.

De plus, si Γ est un réseau uniforme, alors la mesure de Liouville m_{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 μ 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 ω -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 ω -limite pour le flot géodésique est non vide. Montrer, en utilisant le théorème ergodique de Birkhoff 1.7, que si $\mu(R) = 0$, alors μ est une mesure de Lebesgue sur une orbite périodique du flot horocyclique.
- (4) Montrer que si $\mu(R) = 1$, alors μ est la mesure de Liouville, renormalisée pour être de probabilité. Pour cela, on pourra procéder de la manière suivante. ¹⁰²

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 (25) 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 4.15) $\Phi : G \rightarrow T^1\mathbb{H}_{\mathbb{R}}^2$ 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

$$\lim_{n \rightarrow +\infty} M_{t_n} f \rightarrow \frac{1}{\|m_{\text{Liou}}\|} \int_{T^1M} f dm_{\text{Liou}}.$$

102. 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

- 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.

En utilisant le théorème ergodique de Birkhoff 1.7, en déduire que $\mu = \frac{m_{\text{Liou}}}{\|m_{\text{Liou}}\|}$.

(5) Conclure.

4.7 Indications pour la résolution des exercices

Correction de l'exercice E.18. (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 (36) 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 (29), 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 $g \cdot v = (g \cdot \bar{c}(0), i g'(\bar{c}(0)) \cdot \bar{c}'(0)) = (g \cdot \bar{c}(0), i(g \cdot \bar{c})'(0))$. Donc

$$\bar{\mathfrak{h}}^s(g \cdot v) = (g \cdot \bar{c}(s), i(g \cdot \bar{c})'(s)) = (g \cdot \bar{c}(s), i g'(\bar{c}(s)) \cdot \bar{c}'(s)) = g \cdot \bar{\mathfrak{h}}^s(v).$$

Montrons que pour tous les $g \in G$ et $s \in \mathbb{R}$, nous avons

$$\bar{\mathfrak{h}}^s(\Phi(g)) = \Phi(g u_s^-) \quad (37)$$

Puisque $\Phi : G \rightarrow T^1 \mathbb{H}_{\mathbb{R}}^2$ est un C^∞ -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 4.15), ceci montrera que $(\bar{\mathfrak{h}}^s)_{s \in \mathbb{R}}$ est un groupe à un paramètre de C^∞ -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 4.18). 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 (37), 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 (29), 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

$$\forall s, t \in \mathbb{R}, \quad \mathfrak{g}^t \circ \bar{\mathfrak{h}}^s \circ \mathfrak{g}^{-t} = \bar{\mathfrak{h}}^{s e^t} \quad (38)$$

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 Γ est un réseau se démontre comme le corollaire 4.19.

(3) Soient Γ 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 (36) 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.¹⁰³ Avec le même argument (en prenant $t \rightarrow -\infty$) pour le flot horocyclique instable (en utilisant la formule (38)), 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 4.19 et par la formule (37), 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 4.3. Par conséquent, f est constante presque partout, et le flot géodésique est ergodique.

Correction de l'exercice E.19. (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, γ) 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 (28). 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 ω , alors le groupe Γ 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 Γi et $\Gamma \omega$. Puisque les stabilisateurs de i et ω dans Γ sont des groupes cycliques d'ordre 2 et 3 respectivement, ceci démontre la première assertion de la question (2).

103. Cet argument est une variante du phénomène de Mautner décrit dans la partie 4.3.

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) Les géodésiques périodiques de M sont les images par la projection canonique $\mathbb{H}_{\mathbb{R}}^2 \rightarrow M = \Gamma \backslash \mathbb{H}_{\mathbb{R}}^2$ des géodésiques ℓ de $\mathbb{H}_{\mathbb{R}}^2$ définies sur \mathbb{R} qui sont invariantes (modulo reparamétrage par translation) par un élément non trivial $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ dans Γ .

En particulier, γ fixe les deux points (distincts) à l'infini de ℓ , 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 ∞ sont ceux qui vérifient $c' = 0$, donc $a = d = \pm 1$ puisque $a, d \in \mathbb{Z}$ et $ad - ba = 1$. Ils agissent donc par des translations entières $z \mapsto z + k$ où $k \in \mathbb{Z}$, qui n'ont pas d'autre point fixe que ∞ dans la sphère de Riemann $\mathbb{C} \cup \{\infty\}$.

Donc la géodésique ℓ 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 ℓ . Donc les points à l'infini de ℓ 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 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 Γ , enverrait la géodésique périodique ℓ sur une géodésique périodique qui est une demi-droite verticale, ce que nous avons exclu. Donc ℓ admet pour couple d'extrémités un couple de nombres algébriques quadratiques α, α^σ conjugués de Galois (c'est-à-dire solutions distinctes d'une équation quadratique à coefficients entiers).

Réciproquement, soit α un irrationnel quadratique, et montrons que α est un point fixe d'un élément non trivial γ de Γ , ce qui montrera que la géodésique d'extrémité α et son conjugué de Galois est invariante par γ , donc se projette sur une géodésique périodique de M . Quitte à translater α par un entier, 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 α par une homographie entière, nous pouvons donc supposer que le développement en fraction continue de α est pé-

riodique. Par conséquent, α est solution d'une équation $\alpha = \frac{1}{a_1 + \frac{1}{a_2 + \dots + \frac{1}{a_{2n} + \alpha}}}$, donc

est un point fixe d'une homographie entière non triviale (à coefficients entiers strictement

positifs).

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 Γ . 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 , 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 Γ du flot horocyclique.

Le groupe Γ 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, l'image de ∞ par $\begin{bmatrix} p & s \\ q & r \end{bmatrix}$ est exactement $\frac{p}{q}$. Puisque le flot horocyclique commute avec l'action de Γ , nous pouvons donc supposer que le point à l'infini de \mathcal{H} est ∞ . Mais alors l'élément $u_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ est non trivial, et il préserve toute horosphère centrée en ∞ , 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. 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 rappelée des fractions continues.

(6) Si \mathcal{H} est un horocycle, définissons l'*horoboule (ouverte)* 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 de bord \mathcal{H} (union son point à l'infini) si \mathcal{H} est un cercle tangent à l'axe réel (privé de son point à l'infini).

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 Γ , si $\text{HB}_\infty = \{z \in \mathbb{C} : \text{Im } z \geq 1\}$, alors l'image de HB_∞ par un élément $\gamma \in \Gamma$ est ou bien égale à HB_∞ , ou bien d'intérieur disjoint de HB_∞ . Nous avons vu qu'un élément de $\mathbb{R} \cup \{\infty\}$ est le point à l'infini d'une horoboule γHB_∞ pour $\gamma \in \Gamma$ si et seulement s'il appartient à $\mathbb{Q} \cup \{\infty\}$. Toujours par la description du domaine fondamental de Γ , 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 Γ (car \mathcal{F} privé des points intérieurs à HB_∞ est compact). Donc tout rayon géodésique ρ dans $\mathbb{H}_{\mathbb{R}}^2$ ou bien converge vers le point à l'infini de γHB_∞ 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 ρ est positivement récurrente par le flot géodésique dans T^1M si et seulement si le point à l'infini de ρ n'appartient pas à $\mathbb{Q} \cup \{\infty\}$.

(7) Puisque la mesure de Liouville m_{Liou} sur T^1M est finie, puisque le flot géodésique sur T^1M est ergodique (voir le corollaire 4.20, ainsi que l'exercice E.18 (3)), puisque l'application $\text{Re} \circ \text{pr}_2$ est mesurable bornée, donc intégrable pour m_{Liou} , le théorème ergodique de Birkhoff 1.7 dit que pour m_{Liou} -presque tout $v \in T^1M$, la limite suivante existe et a pour valeur celle indiquée, qui se simplifie en utilisant la formule (30),

$$\begin{aligned} \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \text{Re} \text{ pr}_2(\mathbf{g}^t v) dt &= \frac{1}{\|m_{\text{Liou}}\|} \int_{T^1M} \text{Re} \circ \text{pr}_2 dm_{\text{Liou}} \\ &= \int_{z \in \mathcal{F}} \left(\int_{\theta \in \mathbb{R}/\mathbb{Z}} \text{Re}(e^{2i\pi\theta}) d\theta \right) \frac{dx dy}{y^2} = 0. \end{aligned}$$

Correction de l'exercice E.20. (1) Le fait que le flot horocyclique pour la mesure de Liouville m_{Liou} est ergodique a été vu dans le corollaire 4.20, puisque Γ est un réseau (par la proposition 4.16).

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.19.

(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 1.7 et par dénombrabilité, il existe $v \in {}^c R$ tel que, pour tout $f \in A$, nous avons

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t f(\mathbf{h}^s(v)) dt = \mu(f).$$

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 $\mathbb{Q} \cup \{\infty\}$, et par la question (4) de l'exercice E.19, l'orbite de v par le flot horocyclique est périodique, de période $T > 0$. Nous avons donc, pour tout $f \in A$,

$$\mu(f) = \lim_{n \rightarrow +\infty} \frac{1}{nT} \int_0^{nT} f(\mathbf{h}^s(v)) dt = \frac{1}{T} \int_0^T f(\mathbf{h}^s(v)) dt.$$

Par densité de A dans $C_c^0(T^1M)$, ceci est vrai pour tout $f \in C_c^0(T^1M)$, et donc μ 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 (35), puisque Φ est une isométrie et puisque la distance sur G est invariante à gauche,

$$\begin{aligned} d(\Phi(g), \mathbf{g}^t \Phi(g)) &= d(\Phi(g), \Phi(g a_t)) = d(g, g a_t) = d(e, a_t) \\ &\leq \ln(1 + (e^{\frac{|t|}{2}} - 1)) + \ln(1 + (e^{\frac{|t|}{2}} - 1)) = |t|. \end{aligned}$$

De même (et le calcul pour le flot horocyclique instable est identique)

$$d(\Phi(g), \mathfrak{h}^s \Phi(g)) = d(\Phi(g), \Phi(g u_s)) = d(g, g u_s) = d(e, u_s) \leq 2 \ln(1 + |s|) \leq 2|s| .$$

Si $t \geq 0$, alors par la formule (36) (et le calcul pour le flot horocyclique instable est analogue en utilisant la formule (38))

$$\begin{aligned} d(\mathfrak{g}^t \Phi(g), \mathfrak{g}^t \mathfrak{h}^s \Phi(g)) &= d(\mathfrak{g}^t \Phi(g), \mathfrak{h}^s e^{-t} \mathfrak{g}^t \Phi(g)) = d(\Phi(g a_t), \Phi(g a_t u_s e^{-t})) \\ &= d(g a_t, g a_t u_s e^{-t}) = d(e, u_s e^{-t}) \leq 2|s| e^{-t} . \end{aligned}$$

Le premier point en découle avec $\kappa = 2$.

ii) Soit $f \in C_c^0(T^1 M)$. 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 $v_0 \in T^1 M$. 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 $v \in T^1 M$ 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 $u_s u_r^- a_{t'} = \begin{bmatrix} e^{\frac{t'}{2}}(1+r s) & s e^{-\frac{t'}{2}} \\ r e^{\frac{t'}{2}} & e^{-\frac{t'}{2}} \end{bmatrix}$ 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 φ 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 $T^1 M$ d'intérieur non vide, et $N_t f$ est bien définie. De plus, φ est C^∞ 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^∞ . Donc sur K_v , par l'invariance de dm_{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 $v \in T^1 M$ 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,

$$\begin{aligned}
|M_t f(v) - M_t f(v_0)| &\leq 2\epsilon + |N_t f(v) - N_t f(v_0)| \\
&\leq 2\epsilon + \int_{T^1 M} \left| f \circ \mathbf{g}_{-\ln t} \left(\frac{\mathbb{1}_{K_v}}{m_{\text{Liou}}(K_v)} - \frac{\mathbb{1}_{K_{v_0}}}{m_{\text{Liou}}(K_{v_0})} \right) \right| dm_{\text{Liou}} \\
&\leq 2\epsilon + \|f\|_2 \left\| \frac{\mathbb{1}_{K_v}}{m_{\text{Liou}}(K_v)} - \frac{\mathbb{1}_{K_{v_0}}}{m_{\text{Liou}}(K_{v_0})} \right\|_2.
\end{aligned}$$

Puisque la mesure de la frontière de K_v est nulle, la seconde norme \mathbb{L}^2 (qui est indépendante de t) tend vers 0 quand v tend vers v_0 . Ceci montre l'équicontinuité de la famille $(M_t f)_{t \in \mathbb{R}}$.

iii) Soit $v \in R$. Par la question (6) de l'exercice E.19, le vecteur tangent v est alors positivement récurrent par le flot géodésique. Il existe donc 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 v , et en particulier reste dans un compact K de $T^1 M$. Par le théorème d'Arzela-Ascoli, quitte à extraire, il existe une fonction continue $F : T^1 M \rightarrow \mathbb{C}$ telle que nous avons $\lim_{n \rightarrow +\infty} M_{t_n} f \rightarrow F$ pour la convergence uniforme sur K . Donc par la formule (36), nous avons

$$\frac{1}{t_n} \int_0^{t_n} f(\mathbf{h}^s v) ds = \int_0^1 f(\mathbf{h}^{st_n} v) ds = \int_0^1 f(\mathbf{g}^{-\ln t_n} \mathbf{h}^s \mathbf{g}^{\ln t_n} v) ds = M_{t_n} f(\mathbf{g}^{\ln t_n} v) \rightarrow F(v).$$

Par l'ergodicité de la mesure de Liouville (qui donne masse totale à R) par le flot horocyclique, la fonction F est m_{Liou} -presque partout constante. Puisqu'elle est continue, elle est constante, égale à $\frac{1}{\|m_{\text{Liou}}\|} \int_{T^1 M} f dm_{\text{Liou}}$. Pour μ -presque tout $v \in R$, par le théorème ergodique de Birkhoff 1.7, nous avons aussi $\lim_{n \rightarrow +\infty} \frac{1}{t_n} \int_0^{t_n} f(\mathbf{h}^s v) ds = \int_{T^1 M} f d\mu$. Donc $\mu(f) = \frac{1}{\|m_{\text{Liou}}\|} m_{\text{Liou}}(f)$, et par conséquent $\mu = \frac{m_{\text{Liou}}}{\|m_{\text{Liou}}\|}$.

(5) Pour tout $\eta \in \mathbb{R} \cup \{\infty\}$, le flot horocyclique sur $T^1 M$ préserve l'ensemble des vecteurs unitaires tangents v tels que le point à l'infini du rayon géodésique défini par v est égal à η . Donc par la question précédente, l'ensemble mesurable R est invariant par le flot horocyclique. Puisque μ est ergodique, nous avons donc $\mu(R) = 0$ ou $\mu(R) = 1$. Donc les deux questions précédentes concluent.

5 Systèmes dynamiques hyperboliques

Les références principales pour ce chapitre sont [Sma, Bow3, Shu, Yoc2] et nous suivons 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 5.3), un cas local (le théorème de Grobman-Hartman 5.6) et un cas global (le théorème 5.8 de stabilité structurelle d'Anosov¹⁰⁴ 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 5.8).

Outre l'exemple du fer à cheval de Smale¹⁰⁴ traité dans la partie 5.5, 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.

5.1 Endomorphismes linéaires hyperboliques

Soient E un espace de Banach réel et T un endomorphisme linéaire continu¹⁰⁵ 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.¹⁰⁶

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,



Dimitry Anosov
104. 1936-2014



Stephen Smale
1930-

105. 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.

106. 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]).

(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 (κ_s, κ_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 5.1 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 (κ_s', κ_u') à (κ_s, κ_u) dans la définition d'endomorphisme (κ_s, κ_u) -hyperbolique est mineure et purement technique (en particulier utilisé pour la proposition 5.2 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 (κ_s, κ_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

$$\text{ch}(T) = \max\{\|T|_{E^s}\|', \|T|_{E^u}{}^{-1}\|'\} \leq \max\{\kappa_s, \kappa_u^{-1}\} < 1. \quad (39)$$

Lemme 5.1. *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 (κ_s, κ_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

$$E^s = \{x \in E : \forall n \in \mathbb{N}, \quad \|T^n(x)\| \leq c \lambda^n \|x\|\}, \quad (40)$$

dont l'inclusion directe est vérifiée par définition. Une démonstration analogue montre que

$$E^u = \{x \in E : \forall n \in \mathbb{N}, \quad \|x\| \leq c \lambda^n \|T^n(x)\|\},$$

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 (40) 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 $(\frac{\kappa'_s}{\kappa_s})^N, (\frac{\kappa'_u}{\kappa_u})^N \leq \frac{1}{c}$. \square

Rappelons (voir par exemple les rappels de l'appendice 5.11 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 5.1, 5.2 et 5.3, le cas où E est de dimension finie suffit, auquel cas $\text{Sp}(T)$ est l'ensemble des valeurs propres complexes de T .

Proposition 5.2. *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 (κ_s, κ_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 5.11 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 5.2 implique la première.

Supposons que T soit (κ_s, κ_u) -hyperbolique. Alors T n'a pas de valeur propre de module compris entre κ_s et κ_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 λ , alors x_s ou x_u est un vecteur propre non nul de valeur propre λ 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 κ_s et κ_u , alors par la théorie de la réduction¹⁰⁷ des endomorphismes réels en dimension finie, nous avons

107. 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 χ_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 $E = \ker P_s(T) \oplus \ker P_u(T)$.

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 κ_s et strictement supérieur à κ_u . En particulier, l'endomorphisme $T|_{E^u} : E^u \rightarrow E^u$ est inversible, car sans valeur propre nulle. La décomposition de Jordan ¹⁰⁸ de T donne alors l'assertion (3) de la définition des

108. Un bloc de Jordan de taille $d \in \mathbb{N} - \{0\}$ est une matrice $J_d(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix} \in \mathcal{M}_d(\mathbb{C})$,

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_λ 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, 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

$$(\mu I_d + P)^n = \mu^{n-d} \sum_{k=0}^{d-1} \mu^{-k} P^k.$$

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

$$c = \max \left\{ \max_{1 \leq n \leq d-1} \|(\mu I_d + P)^n\|, \frac{\|\sum_{k=0}^{d-1} \mu^{-k} P^k\|}{\kappa^d} \right\}.$$

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 λ_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 $\ell \in \{1, \dots, k\}$, quitte à permuter les blocs de Jordan. Avec les notations de la note de bas de page précédente ¹⁰⁷, nous avons alors

$$E_s = \ker P_s(M) = Q \text{Vect}_{\mathbb{R}}(e_1, \dots, e_{d_1+d_2+\dots+d_\ell}) \text{ et } E_u = \ker P_u(M) = Q \text{Vect}_{\mathbb{R}}(e_{d_1+d_2+\dots+d_\ell+1}, \dots, e_N).$$

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 ^{5.2} en découle.

endomorphismes linéaires hyperboliques avec

$$\kappa'_s \in] \max\{|\lambda| : \lambda \in \text{Sp}(T), |\lambda| < 1\}, 1[\text{ et } \kappa'_u \in] \min\{|\lambda| : \lambda \in \text{Sp}(T), |\lambda| > 1\}, 1[. \quad \square$$

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 1). Nous rappelons que l'espace $C_{ub}^0(E, E)$ 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.¹⁰⁹ Notons

$$\text{Lip}(f) = \sup_{x \neq y} \frac{\|f(x) - f(y)\|}{\|x - y\|}$$

la *constante de Lipschitz* (globale) d'une application $f : E \rightarrow E$. Rappelons qu'une application $f : E \rightarrow E$ est (globalement) δ -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 δ -lipschitzienne, et qu'elle est (globalement) bilipschitzienne si elle est lipschitzienne, bijective, d'inverse lipschitzienne.

Proposition 5.3. (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 δ -lipschitzienne avec $\delta < \delta_0 = \inf\{1 - \text{ch}(T), \|T^{-1}\|^{-1}\}$. 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

$$\|H(0)\| \leq \frac{\|f(0)\|}{\delta_0 - \delta} .$$

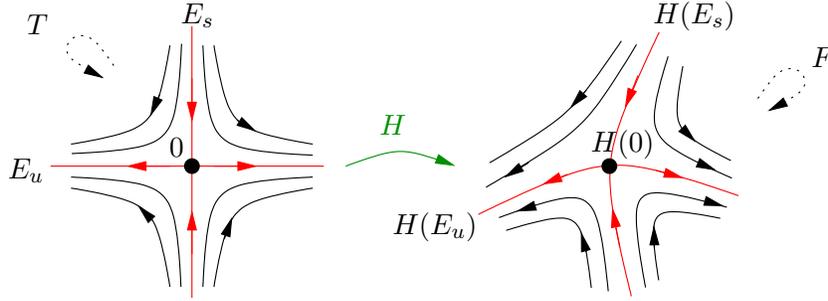
Les images par H de E^s et E^u sont invariantes par F . Privées de $H(0)$, elles sont respectivement contractées et dilatées par F .

109. Rappelons qu'une application $f : X \rightarrow Y$ entre deux espaces métriques est *uniformément continue* si

$$\forall \epsilon > 0, \exists \eta > 0, \forall x, y \in X, \text{ si } d(x, y) < \eta \text{ alors } d(f(x), f(y)) < \epsilon .$$

Le sous-espace vectoriel $C_{ub}^0(E, E)$ 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 $C_{ub}^0(E, E)$ 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

$$\|f(y) - f(x)\| \leq \|f(y) - f_n(y)\| + \|f_n(y) - f_n(x)\| + \|f(x) - f_n(x)\| \leq \epsilon .$$



(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 f est bornée et ϵ -lipschitzienne. Alors $H(0) = 0$ (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 $\delta < \inf\{1 - \text{ch}(T), \|T^{-1}\|^{-1}\}$. 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 δ -lipschitzienne. Montrons que

- (i) il existe un unique élément h de $C_{ub}^0(E, E)$ 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 5.4. Soient $T' \in \mathcal{GL}(E)$ et $g' : E \rightarrow E$ une application δ' -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 φ est contractante sur l'espace métrique complet E . Elle admet donc¹¹⁰ 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

$$\|G'(x) - G'(y)\| \geq \|x - y\| - \|g'(x) - g'(y)\| \geq (1 - \delta') \|x - y\|.$$

110. Voir les notes de bas de page (60) et (61).

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

$$\begin{cases} \tilde{h}_s &= (f_s \circ H + S \circ h_s - g_s) \circ G^{-1} \\ \tilde{h}_u &= U^{-1} \circ (h_u \circ G + g_u - f_u \circ H). \end{cases}$$

Notons que par le lemme 5.4, puisque $\delta < \|T^{-1}\|^{-1}$, l'application G est un homéomorphisme bilipschitzien de E .

Montrons maintenant que l'application $\varphi : C_{ub}^0(E, E) \rightarrow C_{ub}^0(E, E)$ 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 $C_{ub}^0(E, E)$, 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 δ -lipschitziennes et par la formule (39), 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 2.10 (dans la note de bas de page (61)) du point fixe des applications contractantes, appliqué à l'application $(\delta + \text{ch}(T))$ -lipschitzienne φ .

(ii) La majoration de $\|h\|_\infty$ s'obtient à l'aide de la dernière assertion du théorème 2.10, qui dit que

$$d(h, x_0) \leq \frac{1}{1 - (\delta + \text{ch}(T))} d(\varphi(x_0), x_0)$$

pour tout $x_0 \in C_{ub}^0(E, E)$, 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 $h' \in C_{ub}^0(E, E)$ 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 $C_{ub}^0(E, E)$, 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 5.20 d'un meilleur contrôle sur l'unique point fixe p de F (voir la remarque (1) ci-dessus).

Proposition 5.5. *Avec les hypothèses de la proposition 5.3, 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 5.3, 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 δ -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 (39), 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 5.5 découle alors de la dernière assertion du théorème 2.10, car $1 - (\delta + \text{ch}(T)) \geq \delta_0 - \delta$ par la définition de δ_0 . \square

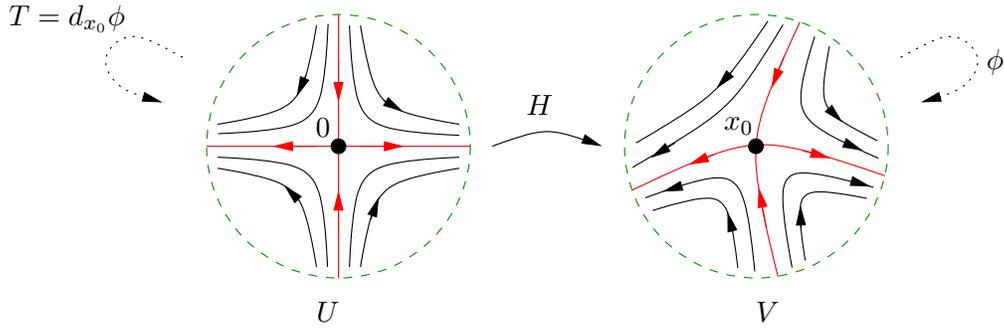
5.2 Le théorème de Grobman-Hartman

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.

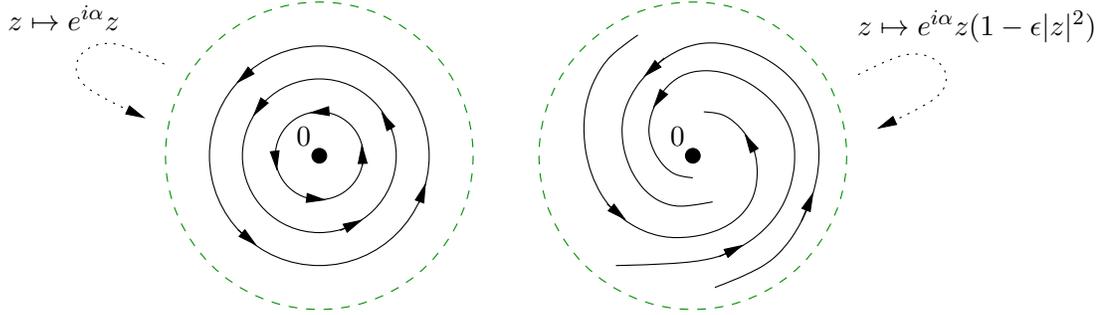
Théorème 5.6. (Théorème de Grobman-Hartman) *Soient $N \in \mathbb{N} - \{0\}$, Ω un ouvert de \mathbb{R}^N , $\phi : \Omega \rightarrow \mathbb{R}^N$ un C^1 -difféomorphisme local et x_0 un point fixe hyperbolique de ϕ (c'est-à-dire x_0 est un point de Ω tel que $\phi(x_0) = x_0$ et tel que la différentielle $T = d_{x_0}\phi$ de ϕ 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 Ω , 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 ω -limites).



Démonstration. Soit $\alpha : \mathbb{R}^N \rightarrow \mathbb{R}$ une fonction plateau de classe C^∞ 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 ϵ assez petit. L'application $F_\epsilon : \mathbb{R}^N \rightarrow \mathbb{R}^N$ coïncide avec ϕ sur $B(0, \epsilon)$ et avec T en dehors de $B(0, 2\epsilon)$. L'application f_ϵ 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

$$d_x f_\epsilon = \alpha\left(\frac{x}{\epsilon}\right) d_x f + \frac{f(x)}{\epsilon} d_{\frac{x}{\epsilon}} \alpha.$$

Notons $M = \max \{ \|\alpha\|_\infty, \|x \mapsto \|d_x \alpha\|_\infty \}$. Nous avons alors

$$\text{Lip}(f_\epsilon) \leq M \left(\sup_{x \in B(0, 2\epsilon)} \|d_x f\| + \sup_{x \in B(0, 2\epsilon)} \frac{\|f(x)\|}{\epsilon} \right) \leq 3M \sup_{x \in B(0, 2\epsilon)} \|d_x f\|$$

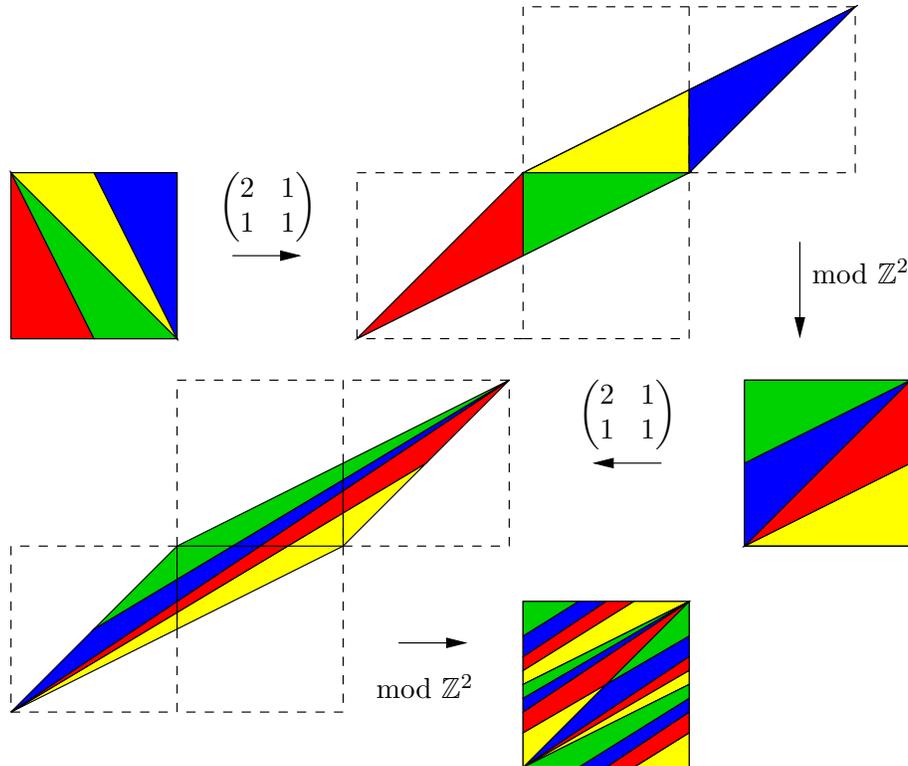
(encore par le théorème des accroissements finis, car $f(0) = 0$). Puisque $d_0 f = 0$, la constante de Lipschitz $\text{Lip}(f_\epsilon)$ tend donc vers 0 quand $\epsilon \rightarrow 0$, et nous pouvons appliquer la proposition 5.3 à F_ϵ pour $\epsilon > 0$ fixé assez petit. Il existe donc un homéomorphisme 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

5.3 Stabilité structurelle des automorphismes hyperboliques du tore

Avant de donner la définition générale d'un système dynamique (uniformément) hyperbolique dans la partie 5.4, 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 3) 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 ϕ_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.

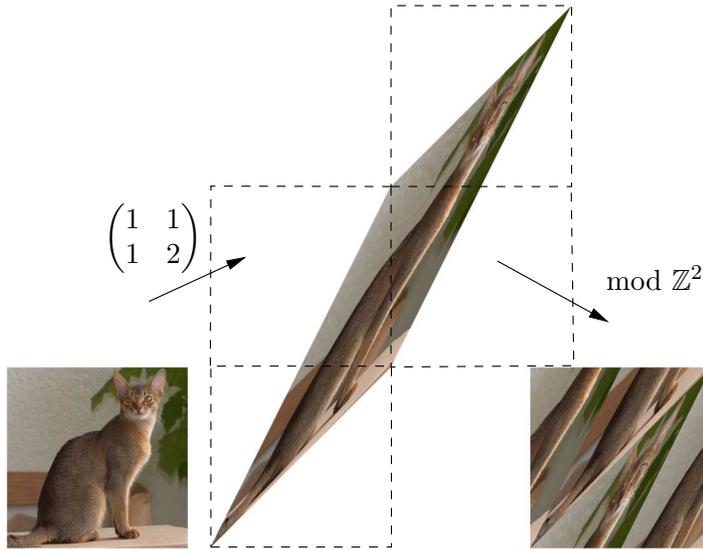


La transformation linéaire de matrice $M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ envoie le carré unité $[0, 1]^2$ sur le parallélogramme de sommets $(0, 0)$, $(2, 1)$, $(3, 2)$ et $(1, 1)$, comme il est facile de vérifier en

regardant l'image des sommets du carrés. Pour comprendre l'application ϕ_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 ϕ_M^2 : nous avons vraiment l'impression que M mélange les couleurs, ce qui illustre le théorème 3.5.

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 5.4.

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 ϕ_M , ce qui donna son nom à l'application ϕ_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 ϕ 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 ϕ_* sur le groupe fondamental \mathbb{Z}^N de \mathbb{T}^N .

Proposition 5.7. 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 ϕ par ρ 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 ϕ tel que $\tilde{\phi}(x_0) = x'_0$.¹¹¹ De plus, si ϕ est un homéomorphisme de \mathbb{T}^N ,

111. Ceci découle du fait que la projection canonique ρ 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

alors $\tilde{\phi}$ est un homéomorphisme de \mathbb{R}^N . **112**

(2) Il existe une matrice $\phi_* \in \mathcal{M}_N(\mathbb{Z})$ à coefficients entiers telle que

$$\forall x \in \mathbb{R}^N, \forall z \in \mathbb{Z}^N, \quad \tilde{\phi}(x+z) = \tilde{\phi}(x) + \phi_*(z). \quad \mathbf{113}$$

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 ϕ est un homéomorphisme de \mathbb{T}^N , alors pour tout $n \in \mathbb{Z}$, nous avons $(\phi^n)_* = (\phi_*)^n$.

(3) La matrice ϕ_* est l'unique matrice $M \in \mathcal{M}_N(\mathbb{Z})$ telle que l'application $\tilde{\phi} - M$ soit bornée sur \mathbb{R}^N . **114**

(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 $\phi_* = \psi_*$. **115** \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 5.7 (1) (et sa note de bas de page 111).

Pour tout point θ 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 θ . C'est, par définition, la *différentielle* de f en θ .

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

$$d(f, g) = \sup_{\theta \in \mathbb{T}^N} (d(f(\theta), g(\theta)) + \|T_\theta f - T_\theta g\|).$$

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.

112. En notant $\tilde{\phi}^{-1}$ l'unique relèvement de l'application continue ϕ^{-1} tel que $\tilde{\phi}^{-1}(x_0') = x_0$, ce qui est possible car $\mathcal{P}(x_0) = \phi^{-1} \circ \phi(\mathcal{P}(x_0)) = \phi^{-1}(\mathcal{P}(x_0'))$, 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 $\tilde{\phi} \circ \tilde{\phi}^{-1} = \text{id}_{\mathbb{R}^N}$.

113. 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 $\phi_*(z) \in \mathbb{Z}^N$. 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 ϕ_* est un morphisme de groupes de \mathbb{Z}^N dans lui-même, et est donc donnée par une matrice à coefficients entiers.

114. 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.

115. Notons $\tilde{\phi}, \tilde{\psi} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ des relèvements de ϕ et ψ . 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.

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 5.16).

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 5.7 (1), nous avons $\widehat{\phi}_M = M$ et $T_\theta \phi_M = M$ pour tout $\theta \in \mathbb{T}^N$. Par l'unicité dans la proposition 5.7 (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.

Théorème 5.8. (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 ϕ_M dans $C^1(\mathbb{T}^N, \mathbb{T}^N)$ tel que toute application ϕ dans \mathcal{U} est topologiquement conjuguée à ϕ_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 5.7 (4), si nous choisissons \mathcal{U} suffisamment petit, tous les éléments ϕ de \mathcal{U} vérifient

$$\phi_* = (\phi_M)_* = M .$$

D'après la proposition 5.7 (3), les relèvements $\tilde{\phi}$ de ϕ s'écrivent $\tilde{\phi} = M + \psi$ avec ψ 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 5.3 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 5.9. *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 5.8, montrons que la conjugaison H entre M et $\tilde{\phi}$ passe au quotient en une conjugaison h entre ϕ_M et ϕ : montrons qu'il existe une application continue $h : \mathbb{T}^N \rightarrow \mathbb{T}^N$ telle que $\rho \circ H = h \circ \rho$. 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 $\rho \circ H' = h' \circ \rho$. Par construction, h' est l'inverse de h . \square

5.4 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.

5.4.1 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 5.10, 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 pr_1, 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 θ est appelée une *trivialisation 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 ϕ* d'un fibré vectoriel E sur K de projection π , à valeurs dans un fibré vectoriel E' sur K' de projection π' , est une application continue $\Phi : E \rightarrow E'$ telle que $\pi' \circ \Phi = \phi \circ \pi$, ou autrement dit le diagramme suivant

$$\begin{array}{ccc} E & \xrightarrow{\Phi} & E' \\ \pi \downarrow & & \downarrow \pi' \\ K & \xrightarrow{\phi} & K' \end{array}$$

est commutatif et, pour tout $x \in K$, l'application $\Phi : E_x \rightarrow E'_{\phi(x)}$ soit linéaire. Un *isomorphisme* 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

$$Tf : v \mapsto \frac{d}{dt}\Big|_{t=0} \phi \circ c$$

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 ϕ .

Soit E un fibré vectoriel sur un espace topologique non vide K , de projection π .

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 π' , 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 ϕ . 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' 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. 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.¹¹⁶ 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

$$\| \Phi \| = \sup_{x \in K} \| \Phi_x \|$$

116. 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 .

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 Φ 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 réel¹¹⁷ 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 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 5.10. *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 d'évaluation en x des sections, définie par $\sigma \mapsto \sigma(x)$ de $\Gamma_b(E)$ dans E_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 σ 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 σ 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

117. 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 $\sigma, \sigma' \in \Gamma^0(E)$ 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

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 θ^{-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 θ^{-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 σ_N et de θ , il existe un voisinage ouvert U_3 de x_0 contenu dans U_2 tel que, pour tout $x \in U_3$, nous avons

$$\|\text{pr}_2 \circ \theta(\sigma_N(x)) - \text{pr}_2 \circ \theta(\sigma_N(x_0))\|_F \leq \frac{\epsilon}{3}.$$

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 σ . \square

5.4.2 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 Λ une partie de M invariante par ϕ , c'est-à-dire telle que $\phi(\Lambda) = \Lambda$. Nous munirons Λ de la topologie induite par M .

Nous noterons encore $\pi : TM|_\Lambda = \bigcup_{x \in \Lambda} T_x M \rightarrow \Lambda$ la restriction de π à la préimage de Λ par π , qui est un fibré vectoriel sur Λ . 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 Λ 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 munirons tout sous-fibré 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 ℓ . Nous supposons que

$$\sup_{x \in \Lambda} \max \{ \|T_x \phi\|_x, \|T_x(\phi^{-1})\|_x \} < +\infty, \quad (41)$$

ce qui automatiquement vérifié si Λ est compact.

Un *champ de vecteurs* sur Λ 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 5.10) des champs de vecteurs X (continus) bornés (ceci est automatique par continuité si Λ est compact) sur Λ pour la norme uniforme

$$\|X\|_\infty = \sup_{x \in \Lambda} \|X(x)\|_x .$$

La transformation ϕ induit un automorphisme linéaire continu ϕ_* de l'espace de Banach $\Gamma_b^0(TM|_\Lambda)$ en posant, pour tout champ de vecteurs $X \in \Gamma_b^0(TM|_\Lambda)$,

$$\phi_*(X) : x \in \Lambda \mapsto T_{\phi^{-1}(x)}\phi(X(\phi^{-1}(x))) \in T_x M , \quad (42)$$

dont l'inverse est $(\phi_*)^{-1} = (\phi^{-1})_*$. Notons que l'application $\phi_*(X)$ est bien continue bornée par la formule (41), 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

$$(\psi \circ \phi)_* = \psi_* \circ \phi_* .$$

Notons $\Gamma_b(TM|_\Lambda)$ l'espace de Banach (voir la proposition 5.10) des champs de vecteurs pas forcément continus sur Λ (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 que ci-dessus, et de l'automorphisme linéaire continu

$$\phi_* : \sigma \mapsto \{ x \mapsto T_{\phi^{-1}(x)}\phi(\sigma(\phi^{-1}(x))) \} \quad (43)$$

défini par la même formule que ci-dessus. En particulier, $\Gamma_b^0(TM|_\Lambda)$ est un sous-espace fermé de $\Gamma_b(TM|_\Lambda)$.

Définition 5.11. *Soit Λ une partie de M invariante par ϕ , munie d'une norme sur $TM|_\Lambda$. Nous dirons que Λ est une partie hyperbolique de M pour ϕ si la propriété (41) est vérifiée et s'il existe $\kappa_s, \kappa_u \in \mathbb{R}$ tels que $0 < \kappa_s < 1 < \kappa_u$, et nous dirons que Λ est une partie (κ_s, κ_u) -hyperbolique de M lorsqu'il convient de préciser, de sorte que l'une des deux conditions équivalentes soit vérifiée :*

- (1) *l'automorphisme linéaire continu ϕ_* de l'espace de Banach $\Gamma_b^0(TM|_\Lambda)$ est (κ_s, κ_u) -hyperbolique,*
- (2) *l'automorphisme linéaire continu ϕ_* de l'espace de Banach $\Gamma_b(TM|_\Lambda)$ est (κ_s, κ_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 Λ continus et $T\phi$ -invariants (c'est-à-dire $T\phi(E^s) = E^s$ et $T\phi(E^u) = E^u$), appelés les sous-fibrés stable et instable de ϕ , 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 5.11 en appendice à ce chapitre 5 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 ϕ_* , le spectre de ϕ_* dans $\Gamma_b^0(TM|_\Lambda)$ est contenu dans le spectre de ϕ_* dans $\Gamma_b(TM|_\Lambda)$. Donc si le spectre de ϕ_* 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 ϕ_* 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 5.2.

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 $X \mapsto X_s + X_u$ où $X_s(x) \in E_x^s$ et $X_u(x) \in E_x^u$ pour tout $x \in \Lambda$, vérifie les axiomes (1) à (3) du début la partie 5.1 définissant le fait que l'automorphisme linéaire continu ϕ_* de $\Gamma_b(TM|_\Lambda)$ (donné par la formule (43)) soit (κ_s, κ_u) -hyperbolique. Donc l'assertion (2) est vérifiée.

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 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 5.12. ¹¹⁸ *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 (κ_s, κ_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 5.1 et le lemme 5.1 (2). Notons $\pi^s : \mathcal{E} \rightarrow \mathcal{E}^s$ et $\pi^u : \mathcal{E} \rightarrow \mathcal{E}^u$ les projections

118. Par souci de complétude, voici une démonstration. L'unicité est immédiate. 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 dans une carte locale en x , puis en multipliant par une fonction plateau pour étendre à tout M , puis en restreignant à Λ). 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 Λ 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 Λ . 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, soit $(\varphi_n)_{n \in \mathbb{N}}$ une suite dans $C^\infty(M)$, dont chaque élément vaille 1 sur un voisinage ouvert de x et soit à valeurs dans $[0, 1]$, telle que le support K_n de φ_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 \mathcal{E} (car $X(x) = 0$). Donc, par le point précédent et par 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 linéarité de A . C'est ce que nous voulions montrer.

linéaires. Remarquons que pour tous les $X \in \mathcal{E}$ et $f \in C_b^0(\Lambda; \mathbb{R})$, par la formule (42) 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 $(\phi_*)^{-n}(f\pi^u(X)) = f \circ \phi^n (\phi_*)^{-n}(\pi^u(X))$. Puisque les applications $f \circ \phi^n$ pour $n \in \mathbb{Z}$ sont uniformément bornées, les propriétés de contraction/dilatation de ϕ_* 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 π^s et π^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

$$\forall X \in \mathcal{E}, \forall x \in \Lambda, \quad (\pi^s X)(x) = p_x^s(X(x)) \quad \text{et} \quad (\pi^u X)(x) = p_x^u(X(x)) .$$

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 de $TM|_\Lambda$ en somme directe, qui vérifient l'assertion (3). \square

Remarques. (1) Si Λ est compact, la définition 5.11 est indépendante du choix de la norme sur $TM|_\Lambda$.

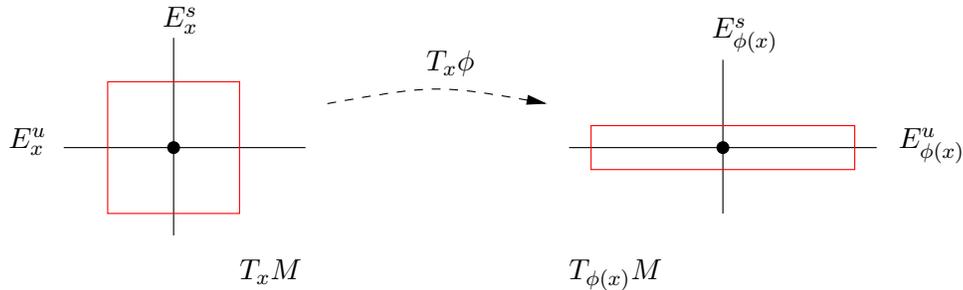
(2) Comme dans le cas linéaire, une décomposition $TM|_\Lambda = E^s \oplus E^u$ en somme directe de deux sous-fibrés, comme dans la définition 5.11, est alors unique. ¹¹⁹

(3) Une norme $\| \cdot \|'$ sur $TM|_\Lambda$ est dite *adaptée* à ϕ si elle est équivalente à la norme $\| \cdot \|$ de $TM|_\Lambda$ (ce qui est automatique si Λ 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

$$\|T\phi|_{E^s}\|' = \sup_{x \in \Lambda} \|T_x\phi|_{E_x^s}\|' \leq \kappa_s, \quad \|(T\phi|_{E^u})^{-1}\|' = \sup_{x \in \Lambda} \|(T_x\phi|_{E_x^u})^{-1}\|' \leq \kappa_u^{-1} .$$



119. En effet, la même démonstration que celle du lemme 5.1 (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 5.11, pour tous les $x \in \Lambda$, 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

$$E_x^u = \{v \in T_x M : \forall n \in \mathbb{N}, \quad \|v\| \leq c\lambda^n \|T_x\phi^n(v)\|\} .$$

Par une démonstration analogue à celle du cas linéaire (voir le lemme 5.1 (3)), il existe au moins une norme adaptée.¹²⁰

(4) Il est possible dans la définition 5.11 (voir par exemple [Yoc2, §1.2]) de remplacer $(\Lambda, TM|_\Lambda)$ par (Λ, E) où Λ est un espace topologique, et $\pi : E \rightarrow \Lambda$ un fibré vectoriel normé au-dessus de Λ , et $(\phi, T\phi)$ par (ϕ, Φ) où $\phi : \Lambda \rightarrow \Lambda$ est un homéomorphisme et $\Phi : E \rightarrow E$ un isomorphisme de fibrés vectoriels au-dessus de ϕ tel que

$$\sup_{x \in \Lambda} \max \{ \|\Phi|_{\pi^{-1}(x)}\|, \|\Phi^{-1}|_{\pi^{-1}(x)}\| \} < +\infty .$$

Voir la sous-partie 5.4.1 pour des définitions.

(5) En fait, la condition de continuité des sous-fibrés stable et instable dans la définition 5.11 est automatique, par le lemme suivant.

Lemme 5.13. *Soit Λ une partie de M invariante par ϕ , munie d'une norme sur le fibré vectoriel $TM|_\Lambda$ vérifiant la propriété (41). 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*

$$\|T_x \phi^n(v)\| \leq c \lambda^n \|v\|, \quad \|v'\| \leq c \lambda^n \|T_x \phi^n v'\|. \quad (44)$$

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 Λ est une partie hyperbolique de M pour ϕ .

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 Λ et TM , convergeant vers $x \in \Lambda$ et $v \in TM$ respectivement. Montrons, par une démonstration analogue¹²¹ à celle du lemme 5.1, 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

$$\begin{aligned} \|v_u\| &= \|T_{\phi^n(x)} \phi^{-n}(T_x \phi^n(v_u))\| \leq c \lambda^n \|T_x \phi^n(v_u)\| = c \lambda^n \|T_x \phi^n(v - v_s)\| \\ &\leq c \lambda^n (\|T_x \phi^n(v)\| + \|T_x \phi^n(v_s)\|) = c \lambda^n \left(\lim_{k \rightarrow +\infty} \|T_{x_k} \phi^n(w_k)\| + \|T_x \phi^n(v_s)\| \right) \\ &\leq c^2 \lambda^{2n} \left(\lim_{k \rightarrow +\infty} \|w_k\| + \|v_s\| \right) = c^2 \lambda^{2n} (\|v\| + \|v_s\|) . \end{aligned}$$

120. 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 (41) que $\sup_{x \in \Lambda} \max \{ \|T_x \phi\|_x, \|T_x(\phi^{-1})\|_x \} < +\infty$ montre que cette norme $\|\cdot\|'$ est équivalente à $\|\cdot\|$.

121. et en fait conséquence de, car la dernière assertion du lemme 5.13 découle aussi du fait que les hypothèses du lemme 5.13 font que la décomposition en somme directe $\Gamma_b(TM|_\Lambda) = \Gamma_b(E^s) \oplus \Gamma_b(E^u)$ vérifie la définition 5.11 (2).

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 $TM|_\Delta$ 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 orthonormée de $T_{x_k}M$, concaténation d'une base 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_xM , 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 $T_xM = E_x^s \oplus E_x^u$, 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) Une orbite périodique *hyperbolique* est 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_xM dans $T_{\phi^p(x)}M = T_xM$, soit un automorphisme linéaire hyperbolique de T_xM . Il est immédiat qu'une orbite périodique hyperbolique est une partie hyperbolique de M pour ϕ , en posant, avec $T_xM = E_x^s \oplus E_x^u$ la décomposition en sous-espace stable et instable de T_xM 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). 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 ϕ .

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 3 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 5.2). 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 5.3).

(3) Soit M une variété différentielle lisse, dont le fibré tangent est muni d'une norme (par exemple riemannienne, indifférente lorsque M est compacte). 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 ϕ est une partie hyperbolique de M pour ϕ ,
- l'ensemble non errant $\Omega(\phi)$ est l'adhérence $\overline{\text{Per}(\phi)}$ de l'ensemble des points périodiques de ϕ .

Exercice E.21. 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). Soit $(\phi^t)_{t \in \mathbb{R}}$ un groupe à un paramètre de classe C^1 de C^1 -difféomorphismes¹²² de M tel que

$$\sup_{x \in M, t \in [-1, 1]} \|T_x \phi^t\|_x < +\infty,$$

ce qui automatiquement vérifié si 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 ϕ , 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 $t \geq 0$,

$$\|T\phi^t|_{E^s}\| \leq c\lambda^t, \quad \|T\phi^{-t}|_{E^u}\| \leq c\lambda^t.$$

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.

Par exemple, si Γ est un sous-groupe discret de $\text{PSL}_2(\mathbb{R})$, alors le flot géodésique $(\mathfrak{g}^t)_{t \in \mathbb{R}}$ sur la variété quotient¹²³ $M = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ 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 Γ 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

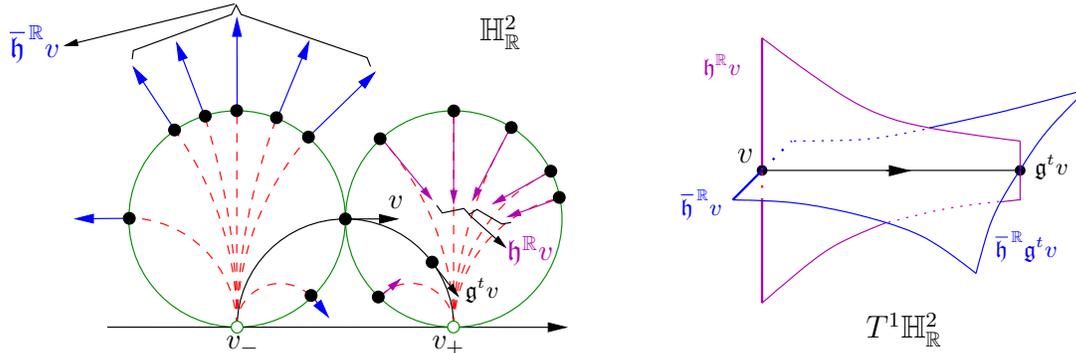
$$\mathfrak{g}^t \circ \mathfrak{g}^\tau = \mathfrak{g}^\tau \circ \mathfrak{g}^t, \quad \mathfrak{g}^t \circ \mathfrak{h}^\tau = \mathfrak{h}^\tau e^{-t} \circ \mathfrak{g}^t \quad \text{et} \quad \mathfrak{g}^t \circ \bar{\mathfrak{h}}^\tau = \bar{\mathfrak{h}}^\tau e^t \circ \mathfrak{g}^t$$

(voir l'exercice E.18 (1) et (2)) et le théorème de dérivation des fonctions composées, nous avons

$$\forall t \in \mathbb{R}, \quad T_x \mathfrak{g}^t(X_v^0) = X_{\mathfrak{g}^t v}^0, \quad T_x \mathfrak{g}^t(X_v^s) = e^{-t} X_{\mathfrak{g}^t v}^s, \quad T_x \mathfrak{g}^t(X_v^u) = e^t X_{\mathfrak{g}^t v}^u.$$

122. 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 .

123. Comme vu dans la partie 4.5, le fibré tangent unitaire $T^1\mathbb{H}_{\mathbb{R}}^2$ s'identifie avec $G = \text{PSL}_2(\mathbb{R})$. Puisque le sous-groupe Γ est discret dans G , l'action par translations à gauche de Γ 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 $T_M = E^0 \oplus E^s \oplus E^u$ 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

$$\forall v \in M, \quad \left\| x_0 X_v^0 + x_s X_v^s + x_u X_v^u \right\|'_v = \max\{|x_0|, |x_s|, |x_u|\}.$$

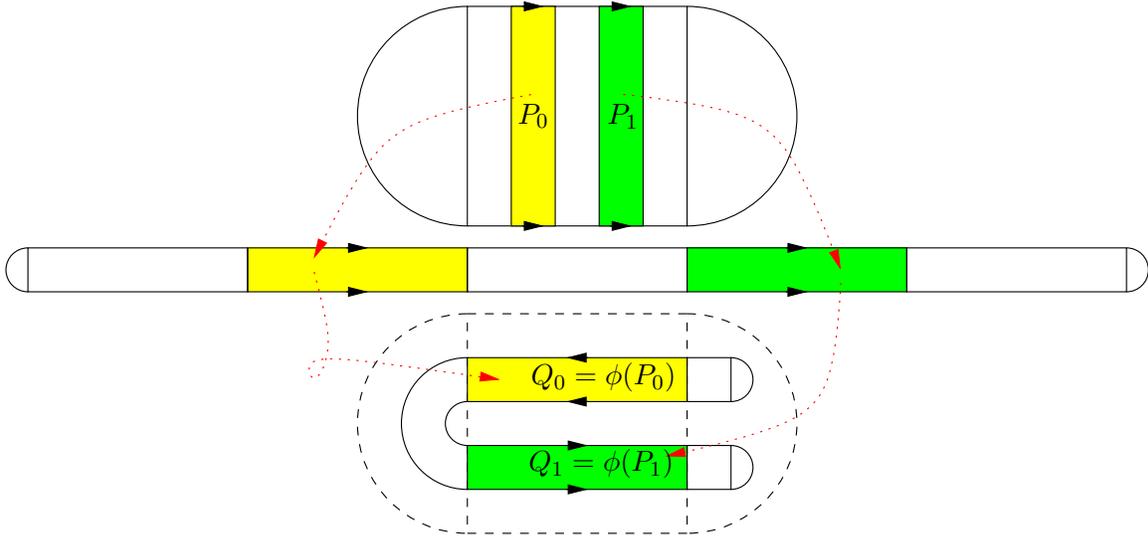
Remarquons que si Γ est trivial, alors l'ensemble non errant de $(\mathbf{g}^t)_{t \in \mathbb{R}}$ est vide. Donc la dynamique du flot géodésique $(\mathbf{g}^t)_{t \in \mathbb{R}}$, même s'il est d'Anosov, ne sera intéressante que lorsque le groupe Γ est intéressant (mais pas forcément de quotient $M = \Gamma \backslash T^1\mathbb{H}_{\mathbb{R}}^2$ compact, comme par exemple pour le groupe modulaire $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, voir la partie 4.5). Nous renvoyons par exemple au joli livre [Dal] pour les aspects de dynamique topologique du flot géodésique sur les surfaces hyperboliques.

5.5 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>

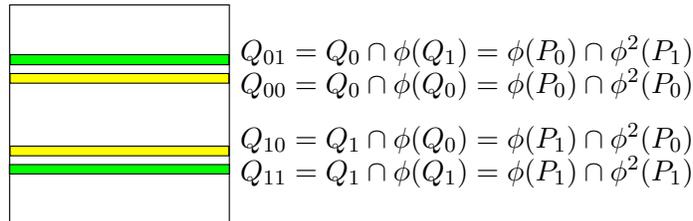
Le *fer à cheval*, introduit par Smale (voir par exemple [Sma]), est l'un des exemples les plus simples de difféomorphisme admettant un ensemble hyperbolique remarquable (en particulier d'intérieur vide, contrairement au cas des difféomorphismes d'Anosov), qui sert d'illustration à de très nombreux phénomènes 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 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 ϕ , 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 π . 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 de manière lisse la transformation ϕ de R dans lui-même en un difféomorphisme encore noté ϕ de la sphère \mathbb{S}_2 , identifiée à $\mathbb{R}^2 \cup \{\infty\}$ par la projection stéréographique, ayant un point fixe répulsif en ∞ .



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

$$Q_w = \bigcap_{i=0}^{n-1} \phi^i(Q_{w_i})$$

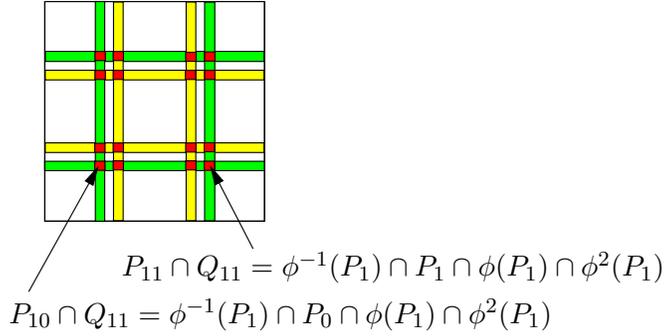
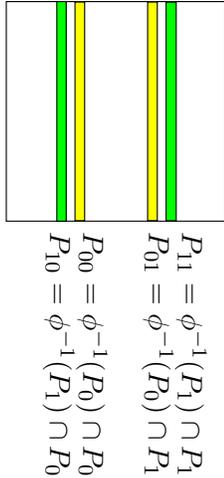
où $w = (w_0, w_1, \dots, w_{n-1}) \in \{0, 1\}^n$. Donc $\Lambda_+ = \bigcap_{i \in \mathbb{N}} \phi^i(Q)$ est un produit $]0, 1[\times K_5$ où K_5 est un espace de Cantor dans $[0, 1]$. De plus,

$$\Lambda_+ = \bigcup_{\omega = (\omega_i)_{i \in \mathbb{N}} \in \{0, 1\}^{\mathbb{N}}} \bigcap_{i \in \mathbb{N}} \phi^i(Q_{\omega_i}),$$

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 ϕ 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

$$P_{00} = \phi^{-1}(P_0) \cap P_0, \quad P_{01} = \phi^{-1}(P_0) \cap P_1, \quad P_{10} = \phi^{-1}(P_1) \cap P_0, \quad P_{11} = \phi^{-1}(P_1) \cap P_1.$$



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 Λ . 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 Λ pour ϕ dans $\mathbb{S}_2 = \mathbb{R}^2 \cup \{\infty\}$. Le *bassin de répulsion* de ∞ pour ϕ , c'est-à-dire le bassin d'attraction de ∞ pour ϕ^{-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 Λ , nous avons (une propriété sur laquelle nous reviendrons plus tard)

$$\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U) = \Lambda . \quad (45)$$

Proposition 5.14. *Le fer à cheval $\phi : \mathbb{S}_2 \rightarrow \mathbb{S}_2$ est un C^∞ -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 ϕ homéomorphe à un espace de Cantor. La restriction de ϕ à \mathbb{R}^2 est un difféomorphisme Axiome A de \mathbb{R}^2 , d'ensemble non errant Λ .*

La restriction de ϕ à Λ 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 ϕ dans Λ est dense dans Λ .

Démonstration. L'espace Λ est le produit $K_5 \times K_5$ de deux espaces de Cantor, donc est un espace de Cantor, et en particulier compact.

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 ϕ à Λ . Par l'exercice E.9, 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 Λ 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 Λ est une partie hyperbolique de \mathbb{S}_2 pour ϕ .

Puisque Λ est un attracteur de bassin d'attraction \mathbb{R}^2 , l'ensemble non errant de la restriction de ϕ à \mathbb{R}^2 est égal à l'ensemble non errant de sa restriction à Λ , donc égal à Λ . Par conséquent, $\phi|_{\mathbb{R}^2}$ est un difféomorphisme Axiome A. \square

5.6 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, Λ une partie de M invariante par ϕ et $\| \cdot \| = (\| \cdot \|_x)_{x \in \Lambda}$ une norme (continue) sur le fibré vectoriel $TM|_\Lambda$ vérifiant la propriété (41) (ce qui est automatique si Λ est compact).

Un *champ de cônes* sur Λ 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

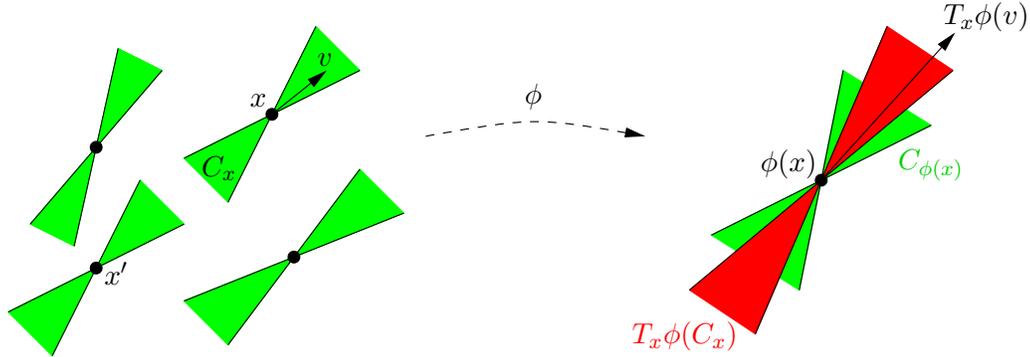
- **(condition de cône)** $C_x = \{v = v_1 + v_2 : v_1 \in E_x^1, v_2 \in E_x^2, \|v_1\|_{1,x} \leq \|v_2\|_{2,x}\}$.
- **(condition de compatibilité avec la norme)** pour tous les $v_1 \in E_x^1$ et $v_2 \in E_x^2$, nous avons $\frac{1}{c} \|v_1 + v_2\|_x \leq \max\{\|v_1\|_{1,x}, \|v_2\|_{2,x}\} \leq c \|v_1 + v_2\|_x$.

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 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 Λ , obtenu en échangeant E_x^1 et E_x^2 .

Notons que si Λ' est une partie de Λ , alors $(C_x)_{x \in \Lambda'}$ est un champ de cônes sur Λ' .



Proposition 5.15. (Critère des champs de cônes instables) Une partie compacte et invariante Λ de M est hyperbolique pour ϕ si et seulement s'il existe un champ de cônes $(C_x)_{x \in \Lambda}$ sur Λ , 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.24 : 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 ϕ^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 Λ soit une partie hyperbolique pour ϕ . Remplaçons la norme de $TM|_\Lambda$ par une norme adaptée. Posons $\lambda_u = \sup_{x \in \Lambda} \|T_x(\phi^{-1})|_{E_x^u}\| < 1$ et $\lambda_s = \sup_{x \in \Lambda} \|T_x\phi|_{E_x^s}\| < 1$, et prenons

$$E_x^1 = E_x^s, \quad E_x^2 = E_x^u, \quad \| \cdot \|_{1,x} = \| \cdot \|_x |_{E_x^s}, \quad \| \cdot \|_{2,x} = \| \cdot \|_x |_{E_x^u},$$

$$\lambda = \frac{1}{\lambda_u \lambda_s}, \quad \mu = \min\left\{\frac{1}{\lambda_u}, \frac{1}{\lambda_s}\right\}, \quad N = 1.$$

Considérons la famille $(C_x)_{x \in \Lambda}$ avec $C_x = \{v_1 + v_2 : v_1 \in E_x^1, v_2 \in E_x^2, \|v_1\|_{1,x} \leq \|v_2\|_{2,x}\}$ pour tout $x \in \Lambda$, qui 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**. Pour tous les $x \in \Lambda$, $v \in C_x$ 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

$$\begin{aligned} \| T_x\phi^N(v) \|_{\phi^N(x)} &= \| (T_x\phi(v))_u \| = \| T_x\phi(v_u) \| \geq \frac{1}{\lambda_u} \| v_u \| = \frac{1}{\lambda_u} \| v \| \\ &\geq \mu \| v \|_x. \end{aligned}$$

De même, puisque $T_x\phi^{-1}(C_x^*) \subset C_{\phi^{-1}(x)}^*$, nous avons

$$\begin{aligned} \| T_x\phi^{-N}(v') \|_{\phi^{-N}(x)} &= \| (T_x\phi^{-1}(v'))_s \| = \| T_x\phi^{-1}(v'_s) \| \geq \frac{1}{\lambda_s} \| v'_s \| = \frac{1}{\lambda_s} \| v' \| \\ &\geq \mu \| v' \|_x. \end{aligned}$$

Réciproquement, soit $(C_x)_{x \in \Lambda}$ un champ de cônes sur Λ vérifiant les propriétés d'inclusion stricte et de dilatation pour ϕ . Montrons qu'il existe, pour tout $x \in \Lambda$, des (uniques) sous-espaces vectoriels E_x^u et E_x^s tels que, avec les notations des remarques précédant l'énoncé de la proposition 5.15,

- (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 5.11, donc Λ sera une partie hyperbolique de M pour ϕ .

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 ϕ .

Cas 1 : Supposons que x n'est pas périodique. Pour tout $y \in \Lambda$, notons $G(y)$ le sous-espace fermé donc compact¹²⁴ 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 $C_{\phi(y)}$. Comme 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 ϕ par ϕ^{-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

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 σ -compacte et à base dénombrable de voisinages, voir par exemple [Pau2, §2.1]), et munie d'une distance induisant sa topologie.

124. 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]

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 métrisable¹²⁵ 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.7 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 5.8 de stabilité structurelle des difféomorphismes Anosov linéaires du tore.

Corollaire 5.16. (Théorème de stabilité des parties hyperboliques) *Soit Λ un compact invariant hyperbolique de M pour ϕ . Alors il existe un voisinage \mathcal{V} de ϕ dans $\text{Diff}^1(M)$ et un voisinage compact U de Λ 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 ψ .*

Démonstration. Nous pouvons supposer que la norme de $TM|_\Lambda$ est adaptée pour ϕ . 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 Λ des sous-espaces vectoriels E_y^1

125. 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.

(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 5.15 pour le difféomorphisme ϕ sont vérifiées en tout point $y \in \Lambda$ puisque la partie Λ est hyperbolique pour ϕ et la norme de $TM|_\Lambda$ est adaptée pour ϕ .

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 ψ 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 Λ_ψ est non vide (sinon en prenant des limites, il existerait des points de Λ dont l'orbite par ϕ sortirait de Λ) et les conditions d'inclusion stricte et de dilatation du champs de cône $(C_y)_{y \in \Lambda_\psi}$ sur l'ensemble Λ_ψ pour le difféomorphisme ψ sont encore vérifiées (quitte à remplacer λ et μ par $\lambda' \in]1, \lambda[$ et $\mu' \in]1, \mu[$). La partie réciproque de la proposition 5.15 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-dessous peut être vide.

Corollaire 5.17. *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

5.7 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 Λ une partie de M invariante par ϕ .

Fixons¹²⁶ une norme $\|\cdot\| : TM \rightarrow \mathbb{R}$ (dite *riemannienne*) euclidienne sur chaque fibre et de carré C^∞ .

126. Pour l'existence des objets $\|\cdot\|$, \exp et d ci-dessous, 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 $x \in \mathcal{M}$, l'application $\exp_x : T_xM \rightarrow M$ est 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.

Fixons un C^∞ -difféomorphisme \exp d'un voisinage de l'image de la section nulle de TM dans un voisinage de la diagonale dans $M \times M$, envoyant la 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 : pour tout compact K de M , il existe $\epsilon_0 > 0$ et $c > 0$ 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

$$\frac{1}{c} \|\exp_{x_0}^{-1}(y) - \exp_{x_0}^{-1}(x)\| \leq d(x, y) \leq c \|\exp_{x_0}^{-1}(y) - \exp_{x_0}^{-1}(x)\|. \quad (46)$$

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 5.18. (Procédure de localisation) *Supposons Λ compact. Il existe $\theta > 0$ assez petit et un voisinage \mathcal{U} assez petit de ϕ 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}_ψ en cas de besoin pour marquer sa dépendance en ψ , 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 ϕ est de classe C^k et si ψ est suffisamment C^k -proche de ϕ , 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 ϕ elle-même et non pas sa perturbation ψ 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 Λ , il existe $\theta_0 > 0$ assez petit et \mathcal{U} un voisinage assez petit de ϕ 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 θ_0 et \mathcal{U} sont assez petits.

Par un argument de recollement standard, nous allons étendre G_x de manière continue en x et 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

$$v \mapsto T_x \phi(v) + \eta\left(\frac{\|v\|^2}{\theta^2}\right) (\exp_{\phi(x)}^{-1} \circ \psi \circ \exp_x(v) - T_x \phi(v)) . \quad (47)$$

Considérons l'application \tilde{G} valant \tilde{G}_x sur chaque fibre $T_x M$ de $TM|_{\Lambda}$. La formule (47), et le fait que le carré de la norme riemannienne soit lisse, montrent que si ϕ et ψ sont de classe C^k sur M , alors il existe un petit voisinage ouvert U de Λ 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 θ_0 et \mathcal{U} sont assez petits, l'application \tilde{G} convient. \square

Le résultat suivant dit en particulier que les orbites par ϕ de deux points distincts d'une partie hyperbolique compacte Λ ne peuvent pas être arbitrairement proches l'une de l'autre pour tous les temps.

Nous dirons qu'un système dynamique topologique (X, ψ) 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 ψ 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 ϵ comme ci-dessus est appelé une *constante d'expansivité* de ψ .

Exemple. Un système de Bernoulli (Σ, σ) 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 5.19. (Expansivité des systèmes dynamiques hyperboliques) *Supposons que Λ soit une partie compacte de M hyperbolique pour ϕ . Il existe un voisinage U de Λ tel que la restriction de ϕ 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 ϕ reste dans U , contient Λ (qui est invariant par ϕ). Donc ce résultat montre en particulier que le C^1 -difféomorphisme ϕ est expansif sur toute partie hyperbolique compacte. Mais en général, $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U)$ peut être strictement plus grand que Λ . Nous reviendrons dans la partie 5.9 sur ce point.

Démonstration. Soit U un voisinage compact de Λ de sorte que $\Lambda' = \bigcap_{n \in \mathbb{Z}} \phi^{-n}(U)$ soit un compact hyperbolique pour ϕ , ce qui est possible par le corollaire de stabilité 5.16.

Notons $\tilde{F} = \tilde{G}_\phi : TM|_{\Lambda'} \rightarrow TM|_{\Lambda'}$ l'application donnée par le lemme 5.18 en prenant $\psi = \phi$ et $\Lambda = \Lambda'$. Si la constante θ 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 δ -lipschitzienne avec δ 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 ϕ_* défini par la formule (43), la différence $f = F - T$ est bornée et δ -lipschitzienne pour la norme uniforme sur $\Gamma_b(TM|_{\Lambda'})$, avec δ arbitrairement petit. Par la proposition 5.3, les endomorphismes linéaires continus T et F sont conjugués.

Soient x et y deux points distincts de Λ' suffisamment proches. Alors il existe un vecteur $v \in T_xM$ non nul tel que $y = \exp_x v$. Soit σ le champ de vecteurs borné (non continu) sur Λ , 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 5.18, ceci implique que l'orbite complète de y ne peut pas rester proche de celle de x . \square

Pour tout $\epsilon > 0$, une ϵ -pseudo-orbite de ϕ 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.24 et qui se démontre par des méthodes analogues, dit que les pseudo-orbites sont proches de vraies orbites.

Théorème 5.20. (Lemme de pistage des systèmes dynamiques hyperboliques)

Supposons que Λ soit une partie compacte de M hyperbolique pour ϕ . Il existe $\epsilon_0 > 0$, $c_0 \geq 1$ et un voisinage U de Λ tel que pour tout $\epsilon \in]0, \epsilon_0]$, pour toute ϵ -pseudo-orbite $(x_n)_{n \in \mathbb{Z}}$ de ϕ 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 ϵ_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 Λ .

Le point x n'appartient pas forcément à Λ , mais il appartient bel et bien à Λ si celui-ci est de plus supposé être localement maximal (voir la partie 5.9).

Démonstration. (Voir par exemple [Yoc2, §2.7].) Soient $\epsilon_0 > 0$ et U un voisinage de Λ , assez petits (à préciser ultérieurement). Soit $(x_n)_{n \in \mathbb{Z}}$ une ϵ -pseudo-orbite de ϕ dans U

avec $\epsilon \leq \epsilon_0$. Pour tout $i \in \mathbb{Z}$, soit y_i l'un des points du fermé Λ 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 $x_i y_i$, pour que x_i soit dans l'image par difféomorphisme local de \exp_{y_i} et y_i dans l'image par difféomorphisme local de $\exp_{\phi(y_{i-1})}$ pour tout $i \in \mathbb{Z}$.

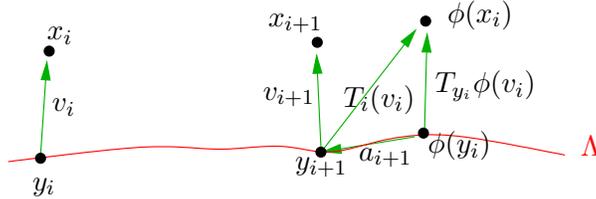
Pour tout $i \in \mathbb{Z}$, soient $E_i = T_{y_i}M$, $E_i^s = E_{y_i}^s$, $E_i^u = E_{y_i}^u$, $v_i \in E_i$ et $a_i \in T_{\phi(y_{i-1})}M$ tels que

$$\exp_{y_i}(v_i) = x_i \quad \text{et} \quad \exp_{\phi(y_{i-1})}(a_i) = y_i .$$

Notons que $\dim E_i^s = \dim E_{i+1}^s$ et $\dim E_i^u = \dim E_{i+1}^u$ si ϵ_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 ϵ_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 5.18, 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

$$\tilde{F}_i = \begin{cases} T_i & \text{sur } \{v \in E_i : \|v\| \geq 2\theta\} \\ \exp_{y_{i+1}}^{-1} \circ \phi \circ \exp_{y_i} & \text{sur } \{v \in E_i : \|v\| \leq \theta\} , \end{cases} \quad (48)$$

et l'application $\tilde{F}_i - T_i$ soit δ -lipschitzienne pour un $\delta > 0$ assez petit (indépendant de i). Si ϵ_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

$$\|v_i\| \leq \frac{\theta}{2} \quad \text{et} \quad \|\tilde{F}_i(v_i) - v_{i+1}\| \leq c_1 \epsilon : \quad (49)$$

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

$$\tilde{F}_i(v_i) = \exp_{y_{i+1}}^{-1} \circ \phi \circ \exp_{y_i}(v_i) = \exp_{y_{i+1}}^{-1}(\phi(x_i)) . \quad (50)$$

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

δ -lipschitzienne avec $\delta > 0$ assez petit. Alors par la proposition 5.3 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 (49) de droite, si $\tilde{v} = (v_i)_{i \in \mathbb{Z}}$, alors $\|F(v) - v\| \leq c_1 \epsilon$ par la définition de la norme de \mathcal{E} . Par la proposition 5.5, 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 ϵ_0 est assez petit, pour tout $i \in \mathbb{N}$, par les formules (48) et (50), 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 (46), 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 .

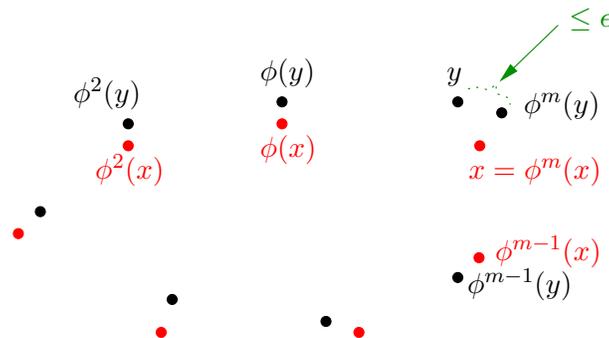
L'unicité de x découle de la propriété d'expansivité (Proposition 5.19), si ϵ_0 et U sont assez petits.

Montrons la dernière affirmation du théorème 5.20. Par le corollaire 5.16, soit V un voisinage compact de Λ tel que $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(V)$ soit un compact hyperbolique pour ϕ . Si ϵ_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 5.21. (Lemme de fermeture d'Anosov) *Supposons que Λ soit une partie compacte de M hyperbolique pour ϕ . Il existe $\epsilon_0 > 0$, $c_0 \geq 1$, et un voisinage U de Λ 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 ϵ_0, c_0, U comme dans l'énoncé du théorème 5.20 et ϵ, m, y comme dans l'énoncé du corollaire 5.21. 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 ϵ -pseudo-orbite de ϕ contenue dans U par les hypothèses du corollaire 5.21. 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 5.20). 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 5.20. \square

Ce lemme de fermeture permet de montrer qu'un système dynamique hyperbolique a de nombreuses orbites périodiques.

Corollaire 5.22. *Supposons que Λ soit une partie compacte de M hyperbolique pour ϕ , et soit V un voisinage compact assez petit de Λ . Alors l'ensemble $\text{Per}(\phi|_{\Lambda'})$ des points périodiques de ϕ dans $\Lambda' = \bigcap_{n \in \mathbb{Z}} \phi^{-n}(V)$ est dense dans l'ensemble non errant $\Omega(\phi|_{\Lambda'})$ de la restriction de ϕ à Λ' .*

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 5.16), si V est assez petit, le compact Λ' est une partie hyperbolique pour ϕ . 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 5.21) pour Λ' . 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 5.21), si ϵ 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 5.20, 5.21 et 5.22 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 lisse¹²⁷ sur M (muni de sa norme $\|\cdot\|$ et de sa distance d choisies au début de la partie 5.7), et $\epsilon, \delta > 0$. Une ϵ -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

$$\left\| \dot{c}(t) - \frac{d}{ds} \Big|_{s=0} \phi^s(c(t)) \right\|_{c(t)} \leq \epsilon.$$

Une courbe lisse $c : \mathbb{R} \rightarrow M$ piste à δ près l'orbite par le flot $(\phi^t)_{t \in \mathbb{R}}$ d'un point $x \in M$ s'il existe un reparamétrage $s : \mathbb{R} \rightarrow \mathbb{R}$ de classe C^1 de c tel que $|\dot{s} - 1| \leq \delta$ et pour tout $t \in \mathbb{R}$, nous ayons

$$d(c \circ s(t), \phi^t(x)) \leq \delta.$$

Théorème 5.23. *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 ϵ -pseudo-orbite (respectivement ϵ -pseudo-orbite fermée) de $(\phi^t)_{t \in \mathbb{R}}$ est pistée à δ près par une orbite (respectivement orbite périodique) de $(\phi^t)_{t \in \mathbb{R}}$. \square*

127. Ceci signifie que l'application $\phi : \mathbb{R} \times M \rightarrow M$ définie par $(t, x) \mapsto \phi^t(x)$ est de classe C^∞ .

5.8 Le théorème des variétés stables et instables

Un problème standard en géométrie différentielle est celui de l'intégrabilité des champs de plans.¹²⁸ Étant donné, en tout point x d'une partie Λ d'une variété différentielle lisse M , un sous-espace vectoriel Δ_x de l'espace tangent $T_x M$, il s'agit de trouver une famille de sous-variétés¹²⁹ de M deux à deux disjointes, dont la réunion contient Λ , et tel que l'espace tangent en tout point x de Λ à la sous-variété passant par x soit exactement Δ_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 Λ 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].

5.8.1 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 5.24. *Soit $F = T + f$ où $f : E \rightarrow E$ est une application δ -lipschitzienne avec $\delta < 1 - \text{ch}(T)$ telle que $f(0) = 0$. Alors il existe une application δ' -lipschitzienne*

128. 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.

129. 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 5.25, ce qui est possible car $1 - \text{ch}(T) < \text{ch}(T)^{-1} - 1$, et en utilisant alors comme constante λ dans l'énoncé 5.24 la valeur

$$\lambda = \delta + \text{ch}(T) < 1.$$

Appelons ensemble κ -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 5.25. Soit $\kappa \in]\text{ch}(T), 1]$. Soit $F = T + f$ où $f : E \rightarrow E$ est une application δ -lipschitzienne avec $\delta < \min\{\kappa - \text{ch}(T), \text{ch}(T)^{-1} - \kappa\}$ telle que $f(0) = 0$. Alors il existe une application δ' -lipschitzienne $G_\kappa : E^s \rightarrow E^u$, avec $\delta' < 1$ et $G_\kappa(0) = 0$, telle que l'ensemble κ -stable de F soit le graphe de l'application G_κ . De plus,

$$\forall x, y \in W_\kappa^s(F), \quad \|F(x) - F(y)\| \leq (\delta + \text{ch}(T)) \|x - y\|, \quad (51)$$

et

$$W_\kappa^s(F) = \{x \in E : \lim_{n \rightarrow +\infty} \kappa^{-n} \|F^n(x)\| = 0\}. \quad (52)$$

En outre, si $\kappa < 1$ et s'il existe $k \geq 1$ tel que f soit de classe C^k , alors l'application G_κ 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 2.10 dans la note de bas de page 61, qui se démontre de la même manière. ¹³⁰

Théorème 5.26. Soient X, Z deux espaces métriques, avec Z complet non vide, $\delta' < 1$ et $\Theta : X \times Z \rightarrow Z$ une application δ' -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 δ' -lipschitzienne.

Si Z est un espace de Banach, si X est un ouvert d'un espace de Banach, et si Θ est de classe C^k pour $k \in (\mathbb{N} - \{0\}) \cup \{\infty\}$, alors φ 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

$$d_x \varphi = (\text{id} - d_{(x, \varphi(x))}^2 \Theta)^{-1} \circ d_{(x, \varphi(x))}^1 \Theta. \quad \square \quad (53)$$

130. Pour tout $x \in X$, l'existence de $\varphi(x)$ est obtenue par application du théorème du point fixe 2.10 dans la note de bas de page 61 à l'application de Z dans Z définie par $z \mapsto \Theta(x, z)$. Le fait que φ est δ' -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 Θ est δ' -lipschitzienne, la norme de l'opérateur linéaire continu $d_{(x, \varphi(x))}^2 \Theta$ de différentielle en $(x, \varphi(x))$ de Θ 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 5.3. 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 δ , 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 $f = F - T$,

$$y_s^{n+1} = F_s(x_s^n + x_u^n) = S(x_s^n) + f_s(x_s^n + x_u^n), \quad (54)$$

$$y_u^n = x_u^n + U^{-1}(x_u^{n+1} - F_u(x_s^n + x_u^n)) = U^{-1}(x_u^{n+1} - f_u(x_s^n + x_u^n)). \quad (55)$$

Notons que si $(y'_s, y'_u) = \Theta(x_s^0, x'_s, x'_u)$, nous avons

$$\|y_s^{n+1} - y'_s{}^{n+1}\| \leq (\|S\| + \delta) \|(x_s^n + x_u^n) - (x'_s{}^n + x'_u{}^n)\|, \quad (56)$$

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 δ' , l'application Θ est δ' -lipschitzienne.

Lemme 5.27. *Soit $k \in (\mathbb{N} - \{0\}) \cup \{+\infty\}$. Si $\kappa < 1$ et si f est de classe C^k , alors Θ 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 Θ dans les formules (54) et (55), et par le fait qu'une application linéaire continue est de classe C^∞ , 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 ℓ de f sont de normes uniformément bornées pour tout $\ell = 1, \dots, k$. En particulier, l'application ℓ -linéaire D_x^ℓ de \mathcal{E}^ℓ 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 ℓ l'application D_x^ℓ pour tout $\ell = 1, \dots, k$. \square

Par le théorème 5.26, 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 (54) et (55) définissant Θ . Donc par l'unicité dans le théorème 5.26, nous avons $\varphi(0) = 0$ et $G_\kappa(0) = 0$. Par le théorème 5.26, l'application φ est δ' -lipschitzienne. Donc G_κ est δ' -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 5.27, l'application Θ est de classe C^k . Donc l'application φ (et par conséquent G_κ) est de classe C^k par le théorème 5.26. De plus, si $d_0 f = 0$, alors par la formule (53) et le fait que $(\Theta(x, 0))_u^0 = U^{-1} f_u(x)$ par la formule (55), nous avons $d_0 G_\kappa = 0$.

Le but d'avoir introduit l'application Θ 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 (55), nous avons l'équivalence

$$\forall n \in \mathbb{N}, F(x_s^n + x_u^n) = x_s^{n+1} + x_u^{n+1} \iff \forall n \in \mathbb{N}, \begin{cases} x_s^{n+1} = y_s^{n+1} \\ x_u^n = y_u^n \end{cases} \quad (57)$$

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_κ .

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_κ est δ' -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 (57) et (56), nous avons

$$\|F_s(x) - F_s(y)\| \leq (\|S\| + \delta) \|x - y\|.$$

Donc

$$\forall x, y \in W_\kappa^s(F), \quad \|F(x) - F(y)\| \leq (\|T|_{E^s}\| + \delta) \|x - y\|. \quad (58)$$

Ceci montre la formule (51). L'assertion (52) du théorème 5.25 en découle, car nous avons $F(0) = 0$ et par l'hypothèse sur δ , nous avons $\|S\| + \delta < \kappa$. \square

5.8.2 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 Λ une partie compacte de M , invariante par ϕ et (κ_s, κ_u) -hyperbolique pour ϕ , où $0 < \kappa_s < 1 < \kappa_u$. Reprenons la norme riemannienne $\|\cdot\|$, l'application riemannienne \exp et la distance riemannienne d de la partie 5.7. Fixons

$$\kappa \in]\max\{\kappa_s, \kappa_u^{-1}\}, 1[. \quad (59)$$

Munissons le fibré vectoriel $TM|_\Lambda$ d'une norme $\|\cdot\|' : TM|_\Lambda \rightarrow \mathbb{R}$ adaptée à ϕ . 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

$$V(z, \epsilon) = \exp_z(\{v \in T_z M : \|v\|' < \epsilon\}). \quad (60)$$

Remarquons que puisque deux normes sur $TM|_\Lambda$ sont équivalentes, et par la formule (46) 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

$$B_d(z, \frac{1}{c'} \epsilon) \subset V(z, \epsilon) \subset B_d(z, c' \epsilon). \quad (61)$$

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 5.28. *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 ¹³¹, 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)

$$\begin{aligned} \forall n \in \mathbb{N}, \quad \forall y, y' \in W_{\text{loc}}^s(x), \quad d(\phi^n(y), \phi^n(y')) &\leq c \kappa^n d(y, y'), \\ \forall n \in \mathbb{N}, \quad \forall y, y' \in W_{\text{loc}}^u(x), \quad d(\phi^{-n}(y), \phi^{-n}(y')) &\leq c \kappa^n d(y, y'), \end{aligned}$$

131. 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 ϵ -voisinage de B et que B soit contenue dans le ϵ -voisinage de A

(4)

$$\begin{aligned} W_{\text{loc}}^s(x) &\text{ est égal à } W_{\epsilon_0}^s(x) = \{y \in M : \forall n \in \mathbb{N}, \phi^n(y) \in V(\phi^n(x), \epsilon_0)\}, \\ W_{\text{loc}}^u(x) &\text{ est égal à } W_{\epsilon_0}^u(x) = \{y \in M : \forall n \in \mathbb{N}, \phi^{-n}(y) \in V(\phi^{-n}(x), \epsilon_0)\}. \end{aligned}$$

Notons que la notation $W_{\text{loc}}^s(x)$ et $W_{\text{loc}}^u(x)$ omet la mention de ϵ , 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 4.5), 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 ϕ par ϕ^{-1} .

Notons $\tilde{F} = \tilde{G}_\phi : TM|_\Lambda \rightarrow TM|_\Lambda$ l'application donnée par le lemme 5.18 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

$$\tilde{F}_x(0) = \exp_{\phi(x)}^{-1} \circ \phi \circ \exp_x(0) = \exp_{\phi(x)}^{-1}(\phi(x)) = 0. \quad (62)$$

Si la constante θ apparaissant dans ce lemme est assez petite, alors pour tout $x \in \Lambda$, l'application $\tilde{F}_x - T_x \phi$ est δ -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 5.25. Rappelons que $\kappa < 1$ a été fixé dans la formule (59).

Lemme 5.29. *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 δ' -lipschitziennes avec $\delta' < 1$, telle que $g_x(0) = 0$ et $d_0 g_x = 0$, et dont le graphe est*

$$\widetilde{W}_\kappa^s(x, \tilde{F}) = \{v \in T_x M : \sup_{n \in \mathbb{N}} \kappa^{-n} \|\tilde{F}^n(v)\|' < +\infty\}. \quad (63)$$

De plus, nous avons

$$\tilde{F}_x(\widetilde{W}_\kappa^s(x, \tilde{F})) \subset \widetilde{W}_\kappa^s(\phi(x), \tilde{F}), \quad (64)$$

$$\forall v, v' \in \widetilde{W}_\kappa^s(x, \tilde{F}), \quad \|\tilde{F}(v) - \tilde{F}(v')\|' \leq (\kappa_s + \delta) \|v - v'\|', \quad (65)$$

$$\forall v \in \widetilde{W}_\kappa^s(x, \tilde{F}), \quad \lim_{n \rightarrow +\infty} \kappa^{-n} \|\tilde{F}^n(v)\|' = 0. \quad (66)$$

Démonstration. Notons $T = \phi_*$ l'application définie par la formule (42) et

$$F : \sigma \mapsto \{x \mapsto \tilde{F}_{\phi^{-1}(x)}(\sigma(\phi^{-1}(x)))\}, \quad (67)$$

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 5.11, 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 ϕ 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 (59). Nous avons $F(0) = 0$ par les formules (67) et (62). Par les explications du début de démonstration du théorème 5.28, l'application $F - T$ est δ -lipschitzienne avec $\delta > 0$ que nous pouvons supposer strictement inférieur à $\min\{\kappa - \text{ch}(T), \text{ch}(T)^{-1} - \kappa\} > 0$.

Puisque ϕ est de classe C^k , par les propriétés de la procédure de localisation (lemme 5.18), 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 5.18, nous avons $d_0 F = T$, donc $d_0(F - T) = 0$ par linéarité de T .

Par le théorème 5.25, il existe donc une application $G_\kappa : \Gamma_b(E^s) \rightarrow \Gamma_b(E^u)$, qui est δ' -lipschitzienne avec $\delta' < 1$ et de classe C^k , telle que $G_\kappa(0) = 0$, $d_0 G_\kappa = 0$ (car $d_0(F - T) = 0$) et par la dernière affirmation du théorème 5.25) et $W_\kappa^s(F)$ soit le graphe de G_κ .

Pour tous les $x \in \Lambda$ et $v \in T_x M$, considérons la section ensembliste bornée σ_v de $TM|_\Lambda$ valant 0 en tout point de Λ différent de x , et valant v en x . Pour tout $x \in M$ fixé, l'application de $T_x M$ dans $\Gamma_0(TM|_\Lambda)$ définie par $v \mapsto \sigma_v$ est linéaire continue, et en particulier, nous avons

$$d_0((\sigma_v)|_{T_x M}) = (\sigma_v)|_{T_x M}.$$

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,

$$v \in \tilde{W}_\kappa^s(\pi(v), \tilde{F}) \iff \sigma_v \in W_\kappa^s(F). \quad (68)$$

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 σ_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 $T_x M$ dans $\Gamma_b(TM|_\Lambda)$, ainsi que l'application de $\Gamma_b(TM|_\Lambda)$ dans $T_x M$ d'évaluation des sections en x , sont linéaires continues, donc de classe C^k . Puisque l'application G_κ 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 $d_0 G_\kappa = 0$, nous avons $d_0 g_x = 0$ 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 67 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_κ . 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 (68). Puisque $W_\kappa^s(F)$ est le graphe de G_κ , 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 δ' -lipschitzienne découle du fait que G_κ 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 5.25 appliqué à T^0 et F^0 , il existe une application G_κ^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 (65) découle du théorème 5.25 (et plus précisément de la formule (51)).

L'anté-pénultième assertion (64) est immédiate par définition des $\widetilde{W}_\kappa^s(x, \widetilde{F})$.

En itérant la formule (65) avec $v' = 0$, la dernière assertion (66) 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 5.29. \square

Revenons à la démonstration du théorème 5.28. Fixons $\epsilon_0 > 0$ assez petit (à préciser ultérieurement), et pour tous les $\epsilon \in]0, \epsilon_0]$ et $x \in \Lambda$, posons

$$W_\epsilon^s(x) = \{y \in M : \forall n \in \mathbb{N}, \phi^n(y) \in V(\phi^n(x), \epsilon)\} , \quad (69)$$

$B_x^s(\epsilon) = \{v \in E_x^s : \|v\|' < \epsilon\}$ 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 5.29. L'application j_x dépend continûment de x . Le second lemme technique suivant décrit les propriétés de cette application j_x .

Lemme 5.30. (i) Si ϵ_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 δ' -lipschitzienne, si ϵ_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 ϵ_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 5.29 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 5.18) et par la définition des voisinages $V(z, \epsilon)$ dans la formule 60, si ϵ_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 (63) et par le fait que \tilde{F} est κ -lipschitzien sur cette intersection infinie par l'assertion (51) du théorème 5.25 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

$$W_\epsilon^s(x) = \exp_x(\tilde{W}_\kappa^s(x, \tilde{F}) \cap \{v \in T_x M : \|v\|' < \epsilon\}). \quad (70)$$

Puisque le graphe de g_x est égal à $\tilde{W}_\kappa^s(x, \tilde{F})$ par le lemme 5.29 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 5.29, l'application \tilde{F} est $(\kappa_s + \delta)$ -lipschitzienne sur $\tilde{W}_\kappa^s(x, \tilde{F})$. \square

Vérifions maintenant les assertions du théorème 5.28 concernant les variétés stables locales. L'assertion (4) est la définition de $W_{\text{loc}}^s(x) = W_{\epsilon_0}^s(x)$.

Le lemme 5.30 (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 5.28 (1) en découle.

L'assertion $\phi(W_{\text{loc}}^s(x)) \subset W_{\text{loc}}^s(\phi(x))$ du théorème 5.28 (2) découle du lemme 5.30 (ii).

Enfin, l'assertion (3) concernant les variétés stables locales découle de la propriété (65) 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 (46) et de l'équivalence des normes. \square

5.8.3 Variétés stables et instables globales des difféomorphismes

Reprenons les notations M, ϕ, Λ 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 ϕ , par

$$W^s(x) = \{y \in M : \lim_{n \rightarrow +\infty} d(\phi^n(x), \phi^n(y)) = 0\}. \quad (71)$$

De même, la *variété instable* (globale) de x , notée $W^u(x)$ ou $W^u(x, \phi)$ lorsqu'il convient de préciser ϕ , est définie par

$$W^u(x) = W^s(x, \phi^{-1}) = \{y \in M : \lim_{n \rightarrow +\infty} d(\phi^{-n}(x), \phi^{-n}(y)) = 0\}. \quad (72)$$

Les variétés stables et instables des points de Λ sont clairement équivariantes par ϕ :

$$\phi(W^s(x)) = W^s(\phi(x)) \quad \text{et} \quad \phi(W^u(x)) = W^u(\phi(x)) . \quad (73)$$

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 $W^s(x) = W^s(y)$.¹³² De même, deux variétés instables sont ou bien disjointes ou bien confondues.

Par la définition des variétés stables et instables, et par le théorème 5.28, nous avons¹³³

$$W^s(x) = \bigcup_{n \in \mathbb{N}} \phi^{-n}(W_{\text{loc}}^s(\phi^n(x))) \quad \text{et} \quad W^u(x) = \bigcup_{n \in \mathbb{N}} \phi^n(W_{\text{loc}}^u(\phi^{-n}(x))) . \quad (74)$$

Par le théorème 5.28 (2), ces unions sont des unions croissantes. Toujours par le théorème 5.28, il existe $\epsilon_2 > 0$ tel que pour tout $x \in \Lambda$, nous ayons

$$W^s(x) = \{y \in M : \limsup_{n \rightarrow +\infty} d(\phi^n(x), \phi^n(y)) \leq \epsilon_2\} ,$$

et il existe $\kappa > 0$ tel que pour tout $x \in \Lambda$, nous ayons

$$W^s(x) = \{y \in M : \lim_{n \rightarrow +\infty} \kappa^{-n} d(\phi^n(x), \phi^n(y)) = 0\} ,$$

et de même pour les variétés instables en remplaçant ϕ par ϕ^{-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.26).

Théorème 5.31. *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 Λ 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 :*

$$T_x(W^s(x)) = E_x^s \quad \text{et} \quad T_x(W^u(x)) = E_x^u .$$

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 5.28. 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 $\exp_{\phi(x)}^{-1} \circ \phi \circ \exp_x$ sur les petits vecteurs de $T_x M$ par le lemme 5.18, puisque le graphe de g_x est égal à $\tilde{W}_\kappa^s(x, \tilde{F})$ (défini par la formule (63)) sur lequel \tilde{F} contracte par le lemme 5.29, il existe $n_0 \in \mathbb{N}$ tel que pour tous les $n \geq n_0$ et $v \in V$, le point

$$\bar{j}_x^s(v) = \phi^{-n} \circ \exp_{\phi^n(x)} \tilde{F}^n(v + g_x(v))$$

ne dépend pas de n . En utilisant la partie gauche de la formule (74) et les propriétés des variétés stables locales, le résultat en découle. \square

132. 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.

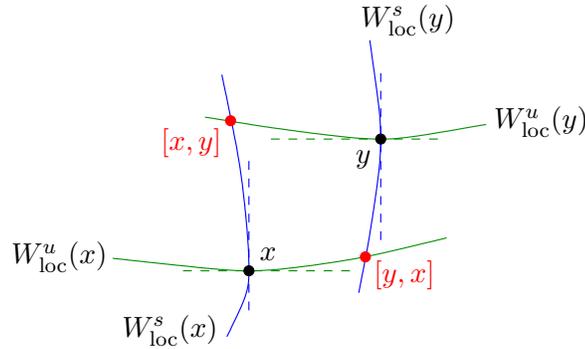
133. L'assertion (4) du théorème 5.28 et la formule (61) montrent que si $y \in W^s(x)$, alors il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$, nous ayons $\phi^n(y) \in V(\phi^n(x), \epsilon_0)$. Donc $\phi^{n_0}(y) \in W_{\text{loc}}^s(\phi^{n_0}(x))$, ce qui donne l'inclusion $W^s(x) \subset \bigcup_{n \in \mathbb{N}} \phi^{-n}(W_{\text{loc}}^s(\phi^n(x)))$. L'assertion (3) du théorème 5.28 montre l'inclusion réciproque.

5.9 La structure de produit local

Reprenons les notations M , ϕ , Λ 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 (69) des variétés stables locales et de la propriété (61) 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 avons

$$[x, y] \in B_d(x, c_1\epsilon) \cap B_d(y, c_1\epsilon) . \quad (75)$$

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 Λ 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 avons $[x, y] \in \Lambda$.

Si Λ 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 Λ est *localement maximale* (ou isolée) s'il existe un voisinage U de Λ tel que $\Lambda = \bigcap_{n \in \mathbb{Z}} \phi^{-n}(U)$.

Théorème 5.32. *Les assertions suivantes sont équivalentes :*

- (1) la partie hyperbolique Λ admet une structure de produit local,
- (2) la partie hyperbolique Λ est localement maximale,
- (3) il existe $\epsilon_2 > 0$ et un voisinage U de Λ tel que pour toute ϵ_2 -pseudo-orbite $(x_i)_{i \in \mathbb{Z}}$ de ϕ 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 5.20 pour Λ), nous avons $x \in \Lambda$.

Démonstration. Montrons que (3) implique (2). Prenons un voisinage U de Λ 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 ϵ -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.

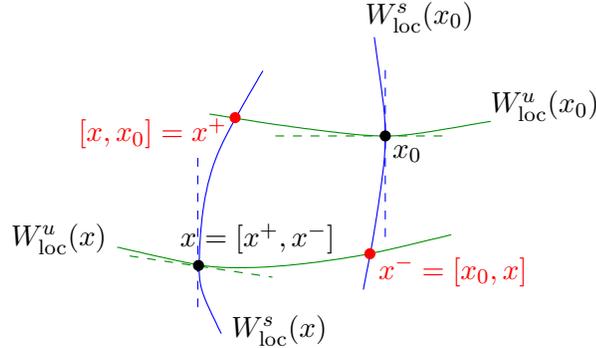
Montrons que (2) implique (1). Soit U un voisinage de Λ tel que $\bigcap_{n \in \mathbb{Z}} \phi^{-n}(U) = \Lambda$. Soit $\epsilon \in]0, \epsilon_0]$ tel que les feuilles locales $W_\epsilon^s(x)$ et $W_\epsilon^u(x)$ soient contenues dans U pour tout $x \in \Lambda$.

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 5.28 (2) des feuilles locales et par l'invariance de Λ par ϕ , 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 Λ , nous pouvons supposer que $x_i \in \Lambda$. Par la dernière affirmation du théorème 5.20, si $\epsilon_2 > 0$ et U sont assez petit, nous pouvons supposer que x appartient à un ensemble hyperbolique compact Λ' contenant Λ .



Si $\epsilon_2 > 0$ est assez petit (devant les constantes nécessaires pour définir le crochet de Bowen pour la partie hyperbolique compacte Λ'), 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 (75), nous avons $x^+, x_- \in B_d(x_0, c_1c_0\epsilon_2)$. Par l'inégalité triangulaire, si $\epsilon_2 > 0$ est assez petit, nous avons donc

$$d(x_-, x_+) \leq 2c_1c_0\epsilon_2 \leq \theta_0(\Lambda)^{-1}\epsilon_1 .$$

Montrons qu'ils appartiennent à Λ , ce qui par l'assertion (1) et la formule centrée ci-dessus implique que x , qui est égal à $[x^+, x^-]$, appartient à Λ . 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 $x_0 \in \Lambda \cap W_{c\epsilon_2}^s(x_0)$ pour tout $c > 0$ et nous posons $x'_0 = x_0$. Supposons x'_i construit. Alors pour ϵ_2 est assez petit devant la constante ϵ_0 du théorème 5.28 (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 (69) des variétés stables locales. Si ϵ_2 est assez petit devant la constante ϵ'_0 de la propriété (61) 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 ϵ_2 est assez petit, puisque la suite $(x_i)_{i \in \mathbb{Z}}$ est une ϵ_2 -pseudo-orbite de ϕ , 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 Λ . 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 Λ étendue par $x'_i = x_i$ pour $i \neq 0$. Si ϵ_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 Λ , comme voulu. \square

Nous dirons qu'une partie Λ d'une variété M munie d'un C^1 -difféomorphisme ϕ est une *pièce basique* de ϕ si elle est invariante, compacte, hyperbolique, localement maximale et si ϕ est transitif sur Λ (c'est-à-dire admet une orbite dense dans Λ).

Par exemple, si $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est le fer à cheval de Smale, alors l'attracteur de Smale Λ est une pièce basique : par construction, le carré unité $Q =]0, 1[^2$ est un voisinage de Λ tel que $\Lambda = \bigcap_{n \in \mathbb{Z}} \phi^n(Q)$, donc Λ est localement maximal. Voir l'exercice E.27 pour une vérification directe que Λ 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 5.33. *Si Λ est une pièce basique d'un C^1 -difféomorphisme ϕ d'une variété M , alors l'ensemble des points périodiques de ϕ dans Λ est dense.*

Démonstration. Ceci découle du corollaire 5.22, 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.10), et du fait qu'une pièce basique qui a un point isolé est réduite à une orbite périodique. \square

5.10 Exercices

Exercice E.22. 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 φ 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.23. 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.24. (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 ϵ -pseudo-orbite de T est ϵ' -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.25. Soient M une variété différentielle lisse et $k \in \mathbb{N} - \{0\}$. Rappelons qu'un point fixe x_0 d'un C^1 -difféomorphisme ϕ 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 .¹³⁴ 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 ϕ un C^1 -difféomorphisme du cercle préservant l'orientation ayant un nombre fini non nul de points fixes, tous hyperboliques.

- Montrer que ϕ admet un nombre pair de points fixes, notés p_1, \dots, p_{2k} dans l'ordre trigonométrique.

- Montrer qu'il existe $\epsilon > 0$ assez petit tel que les arcs ouverts de cercle A_i de longueur 2ϵ centrés en p_i sont deux à deux disjoints, et que si ϕ' est C^1 -proche de ϕ , alors ϕ' a exactement un point fixe p'_i dans chaque arc de cercle de longueur 2ϵ 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 ϕ est structurellement stable.¹³⁵

134. Voir les définitions et propriétés de la topologie C^k avant le corollaire 5.16.

135. Pour tout $k \in (\mathbb{N} - \{0\}) \cup \{\infty\}$, un système dynamique différentiable (X, ϕ) de classe C^k est dit *structurellement stable* s'il existe un voisinage \mathcal{V} de ϕ dans $C^k(X, X)$ pour la topologie C^k tel que tout élément $\phi' \in \mathcal{V}$ soit topologiquement conjugué à ϕ . En général, la conjugaison n'est pas de classe C^k .

(3) Soient $\phi \in \text{Diff}^k(M)$ et U un ouvert de M d'adhérence compacte dans laquelle ϕ admet un et un seul point fixe x_ϕ tels que $x_\phi \in U$ et l'application tangente de ϕ en ce point soit sans valeur propre égale à 1 (respectivement soit hyperbolique). Montrer qu'il existe un voisinage \mathcal{V} de ϕ 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 ϕ' 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 ϕ' est C^1 -proche de ϕ :

- 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 ϕ' admet un point fixe $x_{\phi'}$;
- montrer que ϕ' 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 ϕ' est C^1 -proche de ϕ , 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 ϕ 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 ¹³⁶ 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 Ω 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 Ω , un nombre fini de points fixes, en lesquels la différentielle de ψ 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 ϕ n'a pas la valeur propre 1, est un ouvert dense de $\text{Diff}^k(M)$;
- 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 ϵ tend vers 0.

(5) Montrer qu'il existe un G_δ 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 ϕ est un C^k -difféomorphisme structurellement stable de M , alors ϕ admet, pour tout $n \in \mathbb{N}$, un nombre fini de points périodiques de période n .

Exercice E.26. 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} .

136. 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).

- (1) Montrer que $0 \in \mathbb{T}^2$ est un point fixe hyperbolique de ϕ , 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.27. Soient $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ le fer à cheval de Smale, et Λ l'attracteur de Smale.

- (1) Calculer les points fixes de ϕ . 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 Λ vérifie la propriété de produit local.
- (4) Montrer que ϕ 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.

5.11 Appendice : rappels de théorie spectrale

Nous renvoyons par exemple à [Rud, 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, ¹³⁷ 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).

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 5.34. 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

$$\rho(T) \leq \|T\|.$$

¹³⁷. 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.

(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 holomorphe¹³⁸ 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

$$\forall \lambda \in \mathbb{C}, \quad \forall z \in E_{\mathbb{C}}, \quad \iota(\lambda z) = \bar{\lambda} \iota(z).$$

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 ι , 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 $\|x + iy\|_{\mathbb{C}} \geq \max\{\|x\|, \|y\|\}$.

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 $T_{\mathbb{C}}$, 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 :

$$\|T_{\mathbb{C}}\|_{\mathbb{C}} = \|T\|.$$

138. 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.

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

$$\text{Sp}(T) = \text{Sp}(T_{\mathbb{C}}). \quad (76)$$

Notons que $T_{\mathbb{C}} - \bar{\lambda} \text{id} = \iota \circ (T_{\mathbb{C}} - \lambda \text{id}) \circ \iota$ où ι 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

$$\text{Sp}(T) = \overline{\text{Sp}(T)}. \quad (77)$$

Proposition 5.35. *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 κ'_s n'importe quel élément de $]\rho(T), \kappa_s[$.

Le fait que (3) implique (2) se vérifie comme dans la démonstration du lemme 5.1 (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 5.36. *Soient E un espace de Banach complexe et $T \in \mathcal{L}(E)$. Supposons que $\text{Sp}(T) = \sigma_1 \cup \sigma_2$ avec σ_1 et σ_2 des compacts disjoints, tels qu'il existe une courbe de Jordan lisse $\gamma : \mathbb{S}_1 \rightarrow \mathbb{C}$ contenant σ_1 dans son intérieur et σ_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 γ une courbe de Jordan lisse orientée dans le sens trigonométrique, contenant σ_1 dans la composante connexe bornée de $\mathbb{C} - \gamma$ et σ_2 dans la composante connexe non bornée de $\mathbb{C} - \gamma$. En déformant un peu γ par compacité de σ_2 , soit γ' une courbe de Jordan lisse orientée dans le sens trigonométrique, contenant γ dans la composante connexe

bornée de $\mathbb{C} - \gamma'$ et σ_2 dans la composante connexe non bornée de $\mathbb{C} - \gamma'$. Par le théorème des résidus, puisque γ est dans la composante connexe bornée du complémentaire de γ' , nous avons

$$\forall \lambda' \in \gamma', \quad \frac{1}{2i\pi} \int_{\gamma} \frac{d\lambda}{\lambda - \lambda'} = 0 \quad \text{et} \quad \forall \lambda \in \gamma, \quad \frac{1}{2i\pi} \int_{\gamma'} \frac{d\lambda'}{\lambda - \lambda'} = 1. \quad (78)$$

Puisque γ et γ' sont contenus dans $\mathbb{C} - \text{Sp}(T)$ et par le théorème 5.34 (iv), l'application $\lambda \mapsto R_{\lambda}(T)$ est analytique complexe sur un voisinage de l'anneau de bords γ et γ' . 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 $\lambda, \lambda' \in \mathbb{C} - \text{Sp}(T)$,

$$R_T(\lambda) = R_T(\lambda)(T - \lambda \text{id} + (\lambda - \lambda') \text{id})R_T(\lambda') = R_T(\lambda') + (\lambda - \lambda')R_T(\lambda)R_T(\lambda'). \quad (79)$$

D'où, en utilisant le théorème de Fubini et les formules (79) et (78),

$$\begin{aligned} P_1^2 &= \frac{1}{(2i\pi)^2} \int_{\gamma'} \int_{\gamma} R_T(\lambda)R_T(\lambda') d\lambda d\lambda' = \frac{1}{(2i\pi)^2} \int_{\gamma'} \int_{\gamma} \frac{R_T(\lambda) - R_T(\lambda')}{\lambda - \lambda'} d\lambda d\lambda' \\ &= \frac{1}{2i\pi} \int_{\gamma} R_T(\lambda) \frac{1}{2i\pi} \int_{\gamma'} \frac{d\lambda'}{\lambda - \lambda'} d\lambda + \frac{1}{2i\pi} \int_{\gamma'} R_T(\lambda') \frac{1}{2i\pi} \int_{\gamma} \frac{d\lambda}{\lambda - \lambda'} d\lambda' = P_1. \end{aligned}$$

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

$$\text{Sp}(T) = \text{Sp}(T|_{E_1}) \cup \text{Sp}(T|_{E_2}). \quad (80)$$

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 $P_1 R_T(\lambda) = R_T(\lambda)P_1$. 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 (80) par passage aux complémentaires.]

Notons $\bar{\Omega}$ l'adhérence de la composante connexe bornée de $\mathbb{C} - \gamma$. Elle contient σ_1 dans son intérieur et son complémentaire contient σ_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 (78),

$$(T - \lambda' \text{id})T' = \frac{1}{2i\pi} \int_{\gamma} R_T(\lambda) d\lambda + \int_{\gamma} \frac{d\lambda}{\lambda - \lambda'} \text{id} = P_1 .$$

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 (80), 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 5.37. *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 σ_1 et σ_2 des compacts disjoints de \mathbb{C} , tels qu'il existe une courbe de Jordan lisse $\gamma : \mathbb{S}_1 \rightarrow \mathbb{C}$ contenant σ_1 dans son intérieur et σ_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 5.36, 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 5.36 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 ι 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 (77)) 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 γ et changement de variable $e^{i\theta} \mapsto e^{-i\theta}$ (ce qui est possible par l'hypothèse sur la courbe γ), 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 ι . 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 5.2. Comme déjà dit, la seconde assertion implique la première. Soient $0 < \kappa_s < 1 < \kappa_u$.

Soit $T \in \mathcal{L}(E)$ un élément (κ_s, κ_u) -hyperbolique. Soient E^s et E^u les sous-espaces vectoriels fermés de E comme dans le début de la partie 5.1. 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 (80)), nous avons $\text{Sp}(T) = \text{Sp}(T|_{E^s}) \cup \text{Sp}(T|_{E^u})$. La proposition 5.35 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 σ_1 dans son intérieur et σ_2 dans son extérieur, et est invariante par conjugaison. Donc par le corollaire 5.37, 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 5.35 à $T|_{E_1}$ et à $(T|_{E_2})^{-1}$, le résultat en découle. \square

5.12 Indications pour la résolution des exercices

Correction de l'exercice E.23. Notons P le polynôme $X^4 + 2X^3 + X^2 + 2X + 1$, et

$$M = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \end{pmatrix} \text{ sa matrice compagnon. Le polynôme caractéristique } \det(M - \lambda \text{id})$$

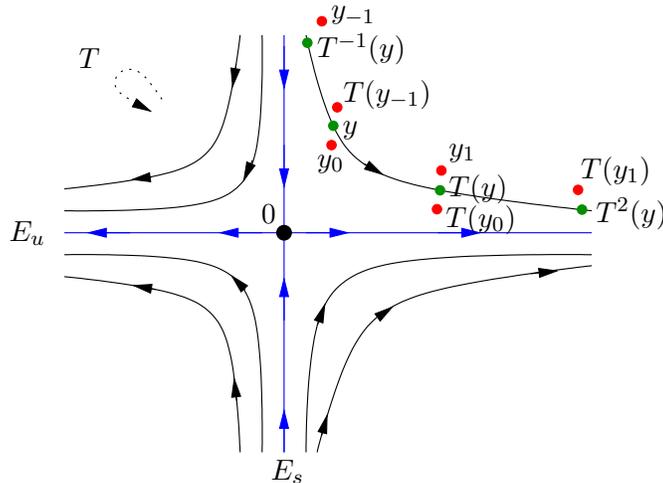
de M est, par récurrence en développant par rapport à la première ligne, égal à $(-1)^4 P$. La matrice M est inversible (car $P(0) \neq 0$) et elle est à coefficients entiers. Les valeurs propres de M sont les racines de P , qui sont

$$\frac{1}{2} \left(-\sqrt{2} - 1 - \sqrt{2\sqrt{2} - 1} \right) \simeq -1,883, \quad \frac{1}{2} \left(-\sqrt{2} - 1 + \sqrt{2\sqrt{2} - 1} \right) \simeq -0,531,$$

$$\frac{1}{2} \left(\sqrt{2} - 1 - i\sqrt{2\sqrt{2} + 1} \right), \quad \frac{1}{2} \left(\sqrt{2} - 1 + i\sqrt{2\sqrt{2} + 1} \right).$$

Il est possible de montrer qu'aucune racine de P n'est une racine de l'unité. Donc la transformation ϕ_M du tore définie par M est mélangeante pour la mesure de Haar du tore, par le théorème 3.5. 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 ϕ 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 5.6.

Correction de l'exercice E.24. 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

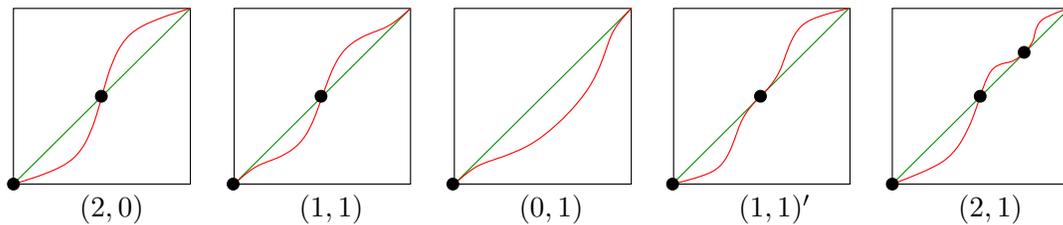
$$y_k - T^k(y) = \sum_{i=0}^{n+k-1} (T^i(y_{k-i}) - T^{i+1}(y_{k-i-1})) + T^{k+n}y_{-n} - T^k y.$$

Donc

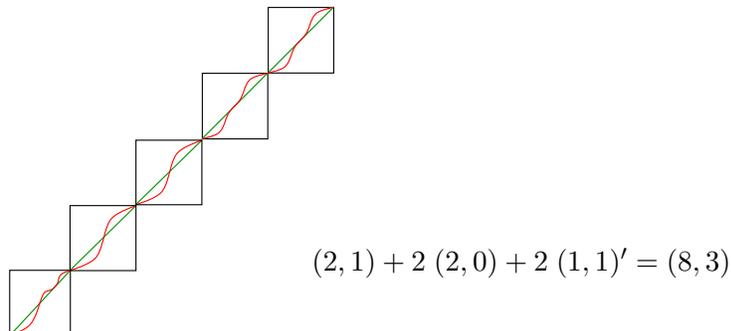
$$\|y_k - T^k(y)\| \leq \sum_{i=0}^{n+k-1} \|T\|^i \|y_{k-i} - T^1(y_{k-i-1})\| + \|T^k\| \|T^n(y_{-n}) - y\|.$$

En faisant tendre n vers $+\infty$, nous avons donc $\|y_k - T^k(y)\| \leq \frac{\epsilon}{1-\|T\|}$.

Correction de l'exercice E.25. (1) Une rotation d'angle non nul (modulo 2π) est un C^1 -difféomorphisme préservant l'orientation sans point fixe. Un C^1 -difféomorphisme ϕ 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 ϕ correspond à un point d'intersection du graphe de ϕ avec la diagonale du carré. Il est hyperbolique si et seulement si la dérivée de ϕ 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 $(2, 1)$.

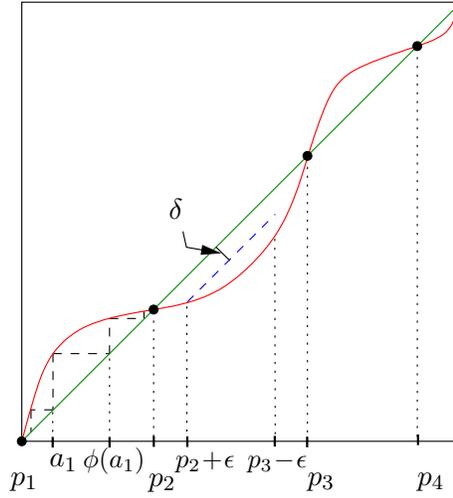


Si $q = 0$ (auquel cas p est supposé pair), on prend $\frac{p}{2}$ blocs $(2, 0)$. Supposons $q \geq 1$. Si $p \leq q$, on prend p blocs $(1, 1)$ et $q - p$ blocs $(0, 1)$. Si $p > q$ et $p - q$ est pair, on prend q blocs $(1, 1)'$ et $\frac{p-q}{2}$ blocs $(2, 0)$. Enfin, si $p > q$ et $p - q$ est impair, 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 ϕ aux points fixes de ϕ 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 ϕ est pair.

• Fixons $\epsilon, \epsilon' > 0$ assez petits tels que les arcs ouverts de cercle A_i de longueur 2ϵ centrés en p_i sont deux à deux disjoints et tels que la dérivée de ϕ sur l'adhérence de A_i soit à distance au moins ϵ' de 1. En dehors de ces arcs, la distance du graphe de ϕ à la diagonale est supérieure à une constante $\delta > 0$. Si ϕ' est C^1 -proche de ϕ , 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 ϕ' à la diagonale est supérieure à $\frac{\delta}{2} > 0$. En particulier, ϕ' n'a pas de point fixe en dehors des A_i .



Par le théorème des valeurs intermédiaires, le graphe de ϕ' 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 ϕ' 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 ϕ 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 ϕ et ϕ' . Ceci montre que ϕ est structurellement stable.

(3) • Pour tout voisinage ouvert U' de x_ϕ d'adhérence compacte et contenue dans U , si ϕ' est uniformément proche de ϕ , alors comme l'application continue $x \mapsto d(\phi(x), x)$ a un minimum δ 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 ϕ' 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 ϕ' est C^1 -proche de ϕ , alors ϕ' 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,¹³⁹ 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 ϕ' admet un point fixe $x_{\phi'}$.

- Montrons que quitte à restreindre U' , l'application ϕ' 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 ϕ , de sorte que ϕ_i admette deux points fixes distincts x_{ϕ_i} et y_{ϕ_i} qui convergent vers 0. Mais alors quitte à extraire, par un développement limité à l'ordre 1 de ϕ_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 ϕ 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 ϕ' est C^1 -proche de ϕ , 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 ψ_y si et seulement si $x \in f^{-1}(y)$. Par le théorème de Sard, qui s'applique en régularité C^1 car Ω 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 ψ_y est C^1 -proche de l'application ψ , 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

139. 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 Ω 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 ϕ 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 ϕ de classe C^k , ayant un nombre fini de points fixes, et en lesquels la différentielle de ϕ n'a pas la valeur propre 1, par une application C^k qui coïncide avec ϕ 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 ϕ difféomorphisme local en 0, fixant 0, et de différentielle en 0 n'ayant pas la valeur propre 1. Pour construire une application φ_ϵ 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^∞ , 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 ϵ est assez petit. Comme $d_0(\phi \circ \varphi_\epsilon) = (1 - \epsilon)d_0\phi$, il suffit donc de prendre ϵ 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 ϕ de M ayant un nombre fini de points périodiques de période n , en lesquels la différentielle de ϕ^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_δ 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 ϕ' est suffisamment C^k -proche de ϕ , alors ϕ'^n est C^k -proche de ϕ^n . Nous renvoyons à [KH] pour la densité.

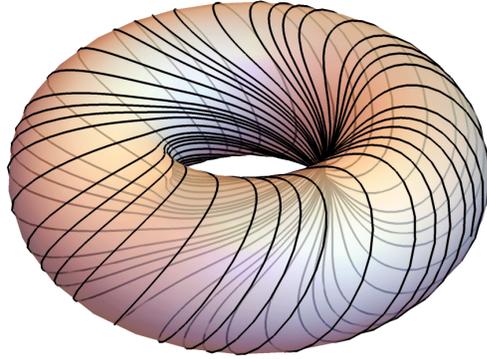
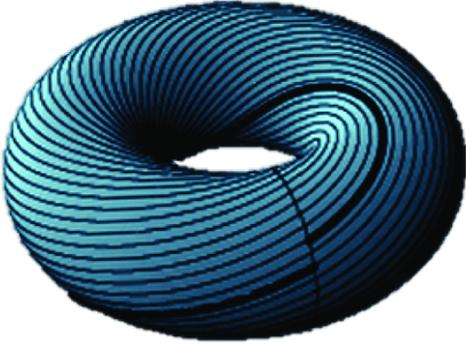
(6) Soit $\phi' \in \mathcal{A}_\infty$ arbitrairement proche de ϕ . Puisque ϕ est structurellement stable, les difféomorphismes ϕ' et ϕ sont conjugués. Puisque qu'une conjugaison induit une bijection de l'ensemble des points fixes de période n de ϕ sur celui de ϕ' , le résultat en découle.

Correction de l'exercice E.26. (1) Il est immédiat que 0 est un point fixe de ϕ 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 ϕ^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 $\mathcal{r} : \mathbb{R}^2 \rightarrow \mathbb{T}^2$ est une isométrie locale, et puisque les variétés stables/instables locales de ϕ sont des parties où ϕ contracte/dilata (voir les formules (3)), les variétés stables et instables locales de ϕ en 0 sont des images par \mathcal{r} 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 (74)), nous avons donc que $W^s(0) = \mathcal{r}(E_0^s)$ et $W^u(0) = \mathcal{r}(E_0^u)$. 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.



6 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 ϵ -près, nous obtenons une fonction de n . L'entropie topologique est, grosso modo, le taux de croissance exponentiel de cette fonction de n lorsque ϵ est très petit.

Le premier but de ce chapitre est de définir précisément l'entropie topologique (partie 6.1) et de montrer (partie 6.3) 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 au cours de Théorie ergodique ou [PY, Wal, CorFS, Pet, EL, Cou] pour tout ce qui concerne l'entropie mesurée des systèmes dynamiques mesurés (mais nous ferons des rappels complets dans la partie 6.2). 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 6.4 pour les systèmes dynamiques symboliques et 6.5 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], [DS, AkS, Tak] et [BrPP, Part II]).

Des références possibles pour ce chapitre sont les livres [Wal, KH, PY] et les notes de cours [BenP].

6.1 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 associative¹⁴⁰ et commutative¹⁴¹. 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 6.1. *Les propriétés suivantes sont satisfaites.*

- i) Nous avons $N(\phi^{-1}\mathcal{U}) \leq N(\mathcal{U})$ avec égalité lorsque ϕ 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 ϕ 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}_ϵ par les boules ouvertes de rayon ϵ 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 ϵ -dense (pour d) si tout point de X est à distance strictement inférieure à ϵ 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 ϵ -séparée (pour d) si deux points distincts de F sont à distance au moins ϵ 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 ϵ -denses de X , et $N'_d(\epsilon)$ la borne supérieure des cardinaux des parties ϵ -séparées de X . Ces deux nombres sont finis, par compacité de X . Nous avons bien sûr $N_d(\epsilon) = N(\mathcal{O}_\epsilon)$.

Lemme 6.2. *Pour tout $\epsilon > 0$, nous avons les inégalités*

$$N'_d(2\epsilon) \leq N_d(\epsilon) \leq N'_d(\epsilon).$$

140. pour tous les $\mathcal{U}, \mathcal{V}, \mathcal{W} \in \mathcal{R}$, nous avons $(\mathcal{U} \vee \mathcal{V}) \vee \mathcal{W} = \mathcal{U} \vee (\mathcal{V} \vee \mathcal{W})$

141. 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 (2ϵ) -séparée n'appartiennent pas à une même boule ouverte de rayon ϵ (par l'inégalité triangulaire). D'autre part, une partie ϵ -séparée de cardinal maximal est forcément ϵ -dense. \square

Définition de l'entropie topologique

Soit (X, ϕ) un système dynamique topologique à temps discret, où X est compact.

Lemme 6.3. *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 6.1 ii) et i), nous avons

$$\begin{aligned} a_{n+m} &\leq \ln N\left(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{U}\right) + \ln N\left(\bigvee_{i=n}^{n+m-1} \phi^{-i}\mathcal{U}\right) \\ &= \ln N\left(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{U}\right) + \ln N\left(\phi^{-n}\left(\bigvee_{i=0}^{m-1} \phi^{-i}\mathcal{U}\right)\right) \leq a_n + a_m. \end{aligned}$$

Le résultat découle alors du lemme classique 2.2 sur les suites sous-additives. \square

La limite donnée par le lemme 6.3 est appelée l'entropie de ϕ 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 ϕ est inversible, par le lemme 6.1 i), nous avons donc

$$h(\phi, \mathcal{U}) = \lim_{n \rightarrow +\infty} \frac{1}{2n+1} \ln N\left(\bigvee_{i=-n}^n \phi^{-i}\mathcal{U}\right) = \inf_{n \in \mathbb{N} - \{0\}} \frac{1}{2n+1} \ln N\left(\bigvee_{i=-n}^n \phi^{-i}\mathcal{U}\right). \quad (81)$$

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 ϕ , 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 6.4. *Soient (X, ϕ) et (X', ϕ') 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

Soit d une distance sur X compatible avec la topologie de X . La proposition suivante fournit une définition plus intuitive de l'entropie topologique de ϕ : elle affirme que nous pouvons nous contenter des recouvrements \mathcal{O}_ϵ .

Notons d_n^ϕ 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 ϵ 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 ϵ près. Nous avons l'égalité

$$N_{d_n^\phi}(\epsilon) = N\left(\bigvee_{i=0}^{n-1} \phi^{-i}\mathcal{O}_\epsilon\right).$$

Proposition 6.5. *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 6.2. \square

Lemme 6.6. (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}_ϵ .*

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 par les 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 des ouverts U de \mathcal{U} .

Corollaire 6.7. *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}_ϵ 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, ϕ) un système dynamique topologique à temps discret, où X est compact.

Proposition 6.8. 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 ϕ 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 ϕ 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 $a_n(\mathcal{U}) = \ln(N(\bigvee_{i=0}^n \phi^{-i} \mathcal{U}))$. Nous avons $a_n(\bigvee_{i=0}^k \phi^{-i} \mathcal{U}) = a_{n+k}(\mathcal{U})$, et lorsque ϕ 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.28. Soit ϕ un homéomorphisme d'un espace métrique compact X . Montrer que $h_{\text{top}}(\phi^{-1}) = h_{\text{top}}(\phi)$.

Remarque. La proposition 6.8 (2) rend naturelle la définition suivante. Soit $\psi = (\psi^t)_{t \in \mathbb{R}}$ un groupe à un paramètre continu¹⁴² d'homéomorphismes d'un espace topologique compact X . L'entropie topologique du système dynamique topologique à temps continu (X, ψ) est définie par

$$h_{\text{top}}(\psi) = h_{\text{top}}(\psi^1).$$

Exercice E.29. 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 ϕ si

$$\lim_{n \rightarrow \infty} \text{diam}\left(\bigvee_{i=0}^n \phi^{-i} \mathcal{U}\right) = 0.$$

Lorsque ϕ est inversible, nous demanderons seulement que

$$\lim_{n \rightarrow \infty} \text{diam}\left(\bigvee_{i=-n}^n \phi^{-i} \mathcal{U}\right) = 0.$$

142. 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 6.9. *Si un recouvrement ouvert \mathcal{U} est générateur pour ϕ , alors*

$$h_{\text{top}}(\phi) = h(\phi, \mathcal{U}) .$$

Démonstration. D'après la proposition 6.8 (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 6.7, cette suite converge vers $h_{\text{top}}(\phi)$. Nous avons donc $h_{\text{top}}(\phi) = h(\phi, \mathcal{U})$. Nous procédons de même lorsque ϕ 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 5.19 pour la définition d'un système dynamique topologique expansif. Notons que le système dynamique (X, ϕ) est expansif si et seulement s'il existe $\epsilon_0 > 0$ (une constante d'expansivité de (X, ϕ)) tel que le recouvrement ouvert \mathcal{O}_{ϵ_0} de X soit générateur pour ϕ .

Corollaire 6.10. *L'entropie topologique $h_{\text{top}}(\phi)$ d'un système dynamique topologique expansif (X, ϕ) , où X est un espace métrique compact, est finie : nous avons*

$$h_{\text{top}}(\phi) < \infty .$$

De plus, si ϵ_0 est une constante d'expansivité de ϕ , 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}_{ϵ_0} soit générateur. Nous utilisons alors la proposition 6.8 (1) et l'égalité précédant la proposition 6.5. \square

6.2 Rappels sur l'entropie mesurée

Soit (X, \mathcal{B}, μ) 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 α , telles que, pour $i \neq j$, nous ayons $A_i \cap A_j = \emptyset$ et $A_1 \cup \dots \cup A_p = X$. Dans la suite de la partie 6.2, les symboles $\alpha, \beta, \gamma, \dots$ désigneront 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 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.

Appelons *partition jointe* de α et β 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} des partitions mesurables de X , qui est associative¹⁴³ et commutative¹⁴⁴. Pour $\alpha_1, \dots, \alpha_n \in \mathcal{P}$, notons

$$\bigvee_{i=1}^n \alpha_i = \alpha_1 \vee \dots \vee \alpha_n = \{A_1 \cap \dots \cap A_n : \forall i = 1, \dots, n, A_i \in \alpha_i, A_1 \cap \dots \cap A_n \neq \emptyset\} .$$

143. pour tous les $\alpha, \beta, \gamma \in \mathcal{P}$, nous avons $(\alpha \vee \beta) \vee \gamma = \alpha \vee (\beta \vee \gamma)$

144. pour tous les $\alpha, \beta \in \mathcal{P}$, nous avons $\alpha \vee \beta = \beta \vee \alpha$

Nous dirons que α est *moins fine* que β (ou que β est *plus fine* que α), et nous noterons $\alpha \preceq \beta$, si tout évènement de α est réunion d'évènements de β . La relation \preceq est une relation d'ordre sur l'ensemble des partitions mesurables de X . Si $\alpha \preceq \beta$ et $\alpha' \preceq \beta'$, alors $(\alpha \vee \alpha') \preceq (\beta \vee \beta')$.

Pour toute partition mesurable α 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 α par ϕ , définie par

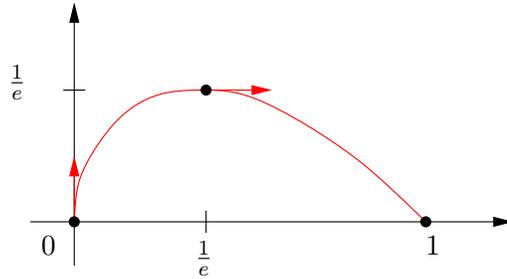
$$\phi^{-1}\alpha = \{\phi^{-1}(A) : A \in \alpha\}.$$

Nous avons $\phi^{-1}(\alpha \vee \beta) = (\phi^{-1}\alpha) \vee (\phi^{-1}\beta)$.

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 concave sur $[0, 1]$, s'annulant en 0 et en 1, de graphe suivant.



L'entropie d'une partition mesurable α est

$$H_\mu(\alpha) = \sum_{A \in \alpha} \varphi(\mu(A)),$$

qui appartient à $[0, +\infty[$ si α est finie, et à $[0, +\infty]$ si α est dénombrable. L'entropie relative d'une partition mesurable α par rapport à une partition mesurable β est

$$H_\mu(\alpha|\beta) = \sum_{A \in \alpha, B \in \beta} \varphi\left(\frac{\mu(A \cap B)}{\mu(B)}\right) \mu(B).$$

Notons qu'en prenant $\beta = \{X\}$ dans cette formule, nous avons

$$H_\mu(\alpha|\{X\}) = H_\mu(\alpha).$$

On note parfois $H(\alpha) = H_\mu(\alpha)$ et $H(\alpha|\beta) = H_\mu(\alpha|\beta)$ lorsque la mesure μ est sous-entendue.

Les entropies et entropies relatives des partitions vérifient les propriétés suivantes (voir le cours de M2 de Théorie ergodique ou [PY, Wal, CorFS, Pet, EL, Cou] ou [BenP, §9]).

Proposition 6.11. *Nous avons les propriétés suivantes.*

- (1) **(Majoration cardinale de l'entropie des partitions)** *Pour toute partition mesurable α de X , nous avons l'inégalité $H_\mu(\alpha) \leq \ln \text{Card}(\alpha)$, avec égalité si et seulement si, pour tout A dans α , nous avons $\mu(A) = \frac{1}{\text{Card}(\alpha)}$.*

- (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 α', β' de X' , nous avons

$$H_\mu(\phi^{-1}\alpha'|\phi^{-1}\beta') = H_{\phi_*\mu}(\alpha'|\beta').$$

En particulier, $H_\mu(\phi^{-1}\alpha') = H_{\phi_*\mu}(\alpha')$.

- (3) **(Monotonie de l'entropie des partitions)** Pour toutes les partitions mesurables α et β de X ,

$$\text{si } \alpha \preceq \beta \text{ alors } H_\mu(\alpha|\gamma) \leq H_\mu(\beta|\gamma),$$

et en particulier $H_\mu(\alpha) \leq H_\mu(\beta)$.

- (4) **(Sous-additivité de l'entropie des partitions)** Pour toutes les partitions mesurables $\alpha, \beta, \gamma, \alpha_1, \dots, \alpha_n$ de X , nous avons

$$H_\mu(\alpha \vee \beta|\gamma) \leq H_\mu(\alpha|\gamma) + H_\mu(\beta|\gamma)$$

et en particulier $H_\mu(\bigvee_{i=1}^n \alpha_i) \leq \sum_{i=1}^n H_\mu(\alpha_i)$.

- (5) **(Concavité de l'entropie des partitions en la mesure)** Soient (X', \mathcal{B}') un espace mesurable et α une partition mesurable de X' . Pour toutes les mesures de probabilités ν_1, \dots, ν_k sur X' et tous les paramètres barycentriques $s_1, \dots, s_k \in [0, 1]$ tels que $\sum_{i=1}^k s_i = 1$, nous avons

$$\sum_{i=1}^k s_i H_{\nu_i}(\alpha) \leq H_{\sum_{i=1}^k s_i \nu_i}(\alpha) \leq \sum_{i=1}^k s_i H_{\nu_i}(\alpha) + \ln k.$$

- (6) **(Continuité de l'entropie des partitions)** Pour toute partition mesurable α de X et pour tout $\eta > 0$, il existe $\epsilon > 0$ tel que, si β est une partition mesurable de X telle que, pour tout A dans α , il existe B dans β avec $B \subset A$ et $\mu(A - B) \leq \epsilon$, alors nous avons $H_\mu(\alpha|\beta) \leq \eta$. \square

Soit $(X, \mathcal{B}, \mu, \phi)$ un système dynamique mesuré, avec μ une mesure de probabilité. Soient α et β des partitions mesurables de X . Puisque ϕ préserve la mesure, par l'assertion (2) ci-dessus, nous avons $H_\mu(\phi^{-1}\alpha|\phi^{-1}\beta) = H_\mu(\alpha|\beta)$, et donc $H_\mu(\phi^{-1}\alpha) = H_\mu(\alpha)$.

Nous appelons *entropie de ϕ relativement à la partition mesurable α* la limite (qui existe, voir la proposition 6.12 (1) ci-dessous)

$$h_\mu(\phi, \alpha) = \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n-1} \phi^{-i}\alpha\right). \quad (82)$$

Il découle de la proposition 6.11 (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 la définition suivante.

Nous appelons *entropie mesurée*¹⁴⁵ (ou parfois juste *entropie*) du système dynamique mesuré $(X, \mathcal{B}, \mu, \phi)$ (ou parfois juste de ϕ lorsque (X, \mathcal{B}, μ) est sous-entendu ou parfois juste de μ lorsque (X, \mathcal{B}, ϕ) est sous-entendu) la quantité

$$h_\mu(\phi) = \sup_{\alpha} h_\mu(\phi, \alpha) \in [0, +\infty],$$

145. Une terminologie plus classique, mais moins appropriée car il n'y a pas de distance impliquée, est celle d'*entropie métrique*.

où la borne supérieure est prise sur toutes les partitions mesurables α de (X, \mathcal{B}, μ) .

Une partition mesurable α de X est dite *génératrice* pour ϕ si la suite des partitions $\alpha_n = \bigvee_{i=0}^n \phi^{-i} \alpha$ engendre la σ -algèbre \mathcal{B} . Lorsque ϕ est inversible, nous demandons seulement que la suite des partitions $\alpha_n = \bigvee_{i=-n}^n \phi^{-i} \alpha$ engendre la σ -algèbre \mathcal{B} .

Nous résumons dans la proposition suivante les propriétés qui nous seront utiles de l'entropie mesurée (voir le cours de M2 de Théorie ergodique ou [PY, Wal, CorFS, Pet, EL, Cou] ou [BenP, §9]).

Proposition 6.12. *Nous avons les propriétés suivantes.*

- (1) **(Existence de l'entropie relative)** *La limite dans la formule (82) 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 ϕ 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) **(Invariance de l'entropie par conjugaison)** *Si $(X, \mathcal{B}, \mu, \phi)$ et $(X', \mathcal{B}', \mu', \phi')$ sont deux systèmes dynamiques mesurés de probabilité conjugués, alors*

$$h_\mu(\phi) = h_{\mu'}(\phi').$$

- (3) **(Inégalité de Rokhlin)** *Nous avons $h_\mu(\phi, \alpha) - h_\mu(\phi, \beta) \leq H_\mu(\alpha|\beta)$.*
- (4) **(Multiplicativité de l'entropie)** *Pour tout $k \in \mathbb{N}$, nous avons $h_\mu(\phi^k) = k h_\mu(\phi)$, et si ϕ est inversible, alors $h_\mu(\phi^k) = |k| h_\mu(\phi)$ pour tout $k \in \mathbb{Z}$.*
- (5) **(Dépendance affine de l'entropie en la mesure)** *Pour tous les $t \in [0, 1]$ et toutes les mesures de probabilité μ, ν invariantes par ϕ 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*

$$h_{t\mu+(1-t)\nu}(\phi) = t h_\mu(\phi) + (1-t) h_\nu(\phi).$$

- (6) **(Théorème de Kolmogorov-Sinai)** *Si une partition mesurable α est génératrice pour ϕ , alors*

$$h_\mu(\phi) = h_\mu(\phi, \alpha). \quad \square$$

Le théorème de Kolmogorov-Sinai est l'un des moyens les plus pratiques de calculer des entropies mesurées, voir les références susdites, ainsi que les exercices E.35 et E.36.

6.3 Principe variationnel

Soit (X, ϕ) un système dynamique topologique, où X est un espace métrique compact. Rappelons (voir la partie 1.2) que nous notons $\text{Prob}_\phi(X)$ l'espace métrisable compact des mesures de probabilité boréliennes sur X invariantes par ϕ , muni de la topologie vague (égale à la topologie faible-étoile).

Théorème 6.13. (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 μ 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 6.14. *Soient X un espace métrisable localement compact séparable, μ 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 $\epsilon = \sup_{1 \leq i \leq p} \mu(A_i - K_i) \geq 0$ 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 6.11 (6), nous pouvons nous arranger pour que $H_\mu(\alpha|\beta) \leq 1$. L'inégalité de Rokhlin (voir la proposition 6.12 (3)) assure que

$$h_\mu(f, \alpha) \leq h_\mu(f, \beta) + 1.$$

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 6.15. *Pour tout $n \in \mathbb{N} - \{0\}$, nous avons $\text{Card} \bigvee_{i=0}^{n-1} f^{-i} \beta \leq 2^n N \left(\bigvee_{i=0}^{n-1} f^{-i} \mathcal{U} \right)$.*

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 6.11 (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

$$h_\mu(\phi, \alpha) \leq h_\mu(\phi, \beta) + 1 \leq h(\phi, \mathcal{U}) + \ln 2 + 1 \leq h_{\text{top}}(\phi) + \ln 2 + 1.$$

Nous en déduisons la majoration $h_\mu(\phi) \leq h_{\text{top}}(\phi) + \ln 2 + 1$. En remplaçant dans cette égalité ϕ par une puissance ϕ^n où $n \in \mathbb{N} - \{0\}$, en utilisant les propositions 6.12 (4) et 6.8 (2), et en divisant par n , nous obtenons $h_\mu(\phi) \leq h_{\text{top}}(\phi) + \frac{\ln 2 + 1}{n}$. 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 6.16. *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 μ 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 6.17. *Soient X un espace métrisable compact, μ une mesure de probabilité borélienne sur X et $\epsilon > 0$. Alors il existe une partition mesurable α de X telle que, pour tout A dans α , nous avons $\text{diam}(A) \leq \epsilon$ et $\mu(\partial A) = 0$.*

Démonstration. Par la finitude de μ , pour tout point x de X , il existe un réel ϵ_x dans $]0, \frac{\epsilon}{2}[$ tel que la sphère de centre x et de rayon ϵ_x dans X soit de mesure nulle pour μ . 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 ϕ de grande entropie mesurée. Notons δ_x la masse de Dirac unité en tout point x d'un espace mesurable.

Lemme 6.18. *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 ϵ -séparée pour la distance d_n^ϕ , 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 μ . De plus, une telle mesure de probabilité μ est invariante par ϕ et nous avons

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln(\text{Card } E_n) \leq h_\mu(\phi).$$

Démonstration. L'existence de μ résulte de la compacité vague et métrisabilité vague de $\text{Prob}(X)$ (voir la proposition 1.5). L'invariance par ϕ de μ résulte de la continuité vague de la mesure image (voir le lemme 1.4) et (sachant que ν_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 6.17, il existe une partition mesurable α de X telle que, pour tout A dans α , 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 α_n contient au plus un point de E_n , chacun d'eux étant de masse $\frac{1}{\text{Card } E_n}$ pour la mesure ν_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 $[x] = \max\{n \in \mathbb{Z} : n \leq x\}$ 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 6.11 (4)), 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

$$\ln(\text{Card } E_n) = H_{\nu_n}(\alpha_n) \leq \sum_{j=0}^{\lfloor \frac{n}{q} \rfloor - 2} H_{\nu_n}(\phi^{-k-aj} \alpha_q) + 3q \ln(\text{Card } \alpha).$$

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 6.11 (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 6.11 (5) pour les paramètres barycentriques constants égaux à $\frac{1}{n+1}$), nous obtenons

$$q \ln(\text{Card } E_n) \leq \sum_{i=0}^n H_{\nu_n}(\phi^{-i} \alpha_q) + 3q^2 \ln(\text{Card } \alpha) \leq (n+1) H_{\mu_n}(\alpha_q) + 3q^2 \ln(\text{Card } \alpha).$$

En divisant par nq et en faisant tendre n vers l'infini, nous obtenons

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln(\text{Card } E_n) \leq \frac{1}{q} \lim_{n \rightarrow \infty} H_{\mu_n}(\alpha_q) = \frac{1}{q} H_\mu(\alpha_q)$$

par le lemme 6.16. Nous faisons alors tendre q vers l'infini, ce qui donne

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln(\text{Card } E_n) \leq h_\mu(\phi, \alpha) \leq h_\mu(\phi).$$

Ceci démontre le lemme 6.18. □

Fin de la démonstration du théorème 6.13. Pour tout $\epsilon > 0$, appliquons le lemme précédent avec E_n de cardinal maximal parmi les parties finies de X qui sont ϵ -séparées pour la distance d_n^ϕ , de sorte que $N'_{d_n^\phi}(\epsilon) = \text{Card } E_n$. Pour tout $\epsilon > 0$, il existe donc un élément μ_ϵ dans $\text{Prob}_\phi(X)$ tel que

$$\sup_{\mu \in \text{Prob}_\phi(X)} h_\mu(\phi) \geq h_{\mu_\epsilon}(\phi) \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \ln N'_{d_n^\phi}(\epsilon).$$

D'après la proposition 6.5, le membre de droite tend vers $h_{\text{top}}(\phi)$ quand ϵ 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 6.19. *Nous avons $h_{\text{top}}(\phi) = h_{\text{top}}(\phi|_{\Omega(\phi)})$*

Démonstration. Par la proposition 1.9, le support de toute mesure finie invariante par ϕ est contenu dans l'ensemble non errant $\Omega(\phi)$ de ϕ . Le résultat découle alors du principe variationnel. □

Soit (X, ϕ) un système dynamique topologique, où X est un espace topologique localement compact. Une *mesure d'entropie maximale* de (X, ϕ) est une mesure de probabilité borélienne μ_{max} sur X invariante par ϕ 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 6.12 (2) et la proposition 6.4, dit que ces problèmes d'existence et d'unicité sont invariants par conjugaison.

Proposition 6.20. *Soient (X, ϕ) et (X', ϕ') 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, ϕ) dans l'ensemble des mesures d'entropie maximale de (X', ϕ') . □*

En particulier, le système dynamique (X, ϕ) 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', ϕ') .

Corollaire 6.21. *Un système dynamique topologique expansif (X, ϕ) , où X est un espace métrique compact, admet au moins une mesure d'entropie maximale μ_{\max} telle que $h_{\mu_{\max}}(\phi) = h_{\text{top}}(\phi)$.*

Démonstration. Par la deuxième affirmation du corollaire 6.10, il existe $\epsilon_0 > 0$ tel que

$$h_{\text{top}}(\phi) = \limsup_{n \rightarrow \infty} \frac{1}{n} \ln N'_{d_n^f}(\epsilon_0).$$

En appliquant le lemme 6.18 avec E_n de cardinal maximal parmi les parties finies de X qui sont ϵ_0 -séparées pour la distance d_n^ϕ , il existe donc un élément μ 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.35).

(2) Si (X, ϕ) est un décalage de Bernoulli sur un alphabet fini \mathcal{A} , si ν 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.36). Nous généraliserons ceci dans la partie 6.4 pour les sous-décalages de type fini.

(3) Nous montrerons dans la partie 6.5 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.

6.4 Entropie topologique des systèmes dynamiques symboliques

Les références recommandées pour cette partie sont [PY, BenP].

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, ϕ) 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 ϕ est inversible) *code* le point x si, à chaque temps n , le point $\phi^n(x)$ appartient à R_{ω_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 ω pour lesquelles cette intersection est non vide. Le système dynamique (X, ϕ) 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

$$(a_n)_{n \in \mathbb{N}} \mapsto \sum_{n \in \mathbb{N}} \frac{a_n}{10^{n+1}}.$$

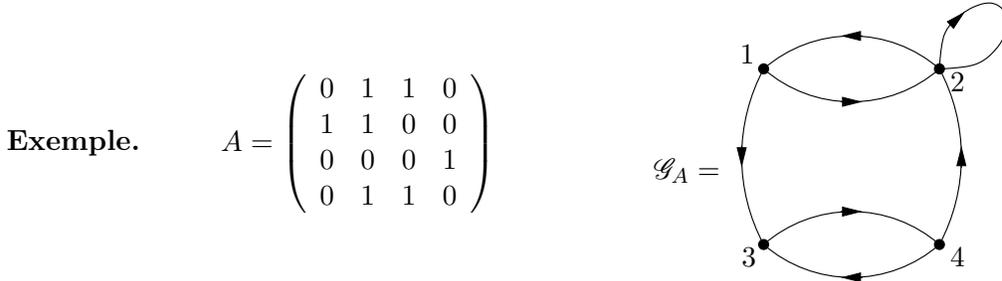
Un autre exemple fondamental de codage est celui que nous avons effectué pour l'étude du fer à cheval sur l'attracteur de Smale (voir la partie 5.5). Dans cet exemple, la semi-conjugaison était une conjugaison, et il n'y avait pas besoin de passer à un sous-système du système de Bernoulli, mais cette 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, Sar3, Sar1, Sar4, Woe].

6.4.1 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*¹⁴⁶ du graphe orienté \mathcal{G}_A .



Nous appellerons *sous-décalage de type fini*¹⁴⁷ bilatère de matrice d'incidence A (ou de graphe \mathcal{G}_A) le système dynamique topologique inversible (à temps discret) (Σ_A, σ_A) où

- Σ_A est la partie fermée donc compacte de $\Sigma = \mathcal{A}^{\mathbb{Z}}$ définie par

$$\Sigma_A = \{ \omega \in \mathcal{A}^{\mathbb{Z}} : \forall n \in \mathbb{Z}, A_{\omega_n, \omega_{n+1}} = 1 \},$$

- σ_A est la restriction du décalage de Bernoulli bilatère σ à Σ_A .

Nous munissons Σ_A de la distance induite d et de la topologie induite par celles de Σ (voir la formule (4)).

Nous définissons de même le *sous-décalage de type fini*¹⁴⁸ 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 (Σ, σ) par (Σ_+, σ_+)) 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 $(\Sigma_{A,+}, \sigma_{A,+}) = (\Sigma_+, \sigma_+)$.

Une matrice carrée $B = (b_j)_{i,j \in I}$ à coefficients positifs ou nuls est dite

- *irréductible* si pour tous les i, j dans I , il existe $n > 0$ tel que $(A^n)_{ij} \neq 0$,

146. Certains ouvrages, dont [Kit], disent *matrice de transition*, mais nous préférons réserver cette terminologie aux matrices de transition des chaînes de Markov, voir la sous-partie suivante.

147. 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.

148. Voir la note de bas de page précédente.

• *apériodique*¹⁴⁹ s'il existe un entier $k \in \mathbb{N} - \{0\}$ tel que les coefficients de B^k soient tous strictement positifs.¹⁵⁰

La condition d'apériodicité est plus forte que celle d'irréductibilité.¹⁵¹ Voir les exercices E.30 ci-dessous et E.37.

Exercice E.30. 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 (Σ_A, σ_A) est positivement transitif,
- (4) A n'a pas de ligne ou de colonne nulle et le système dynamique topologique inversible (Σ_A, σ_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 ω -limite est égal à $\Sigma_{A,+}$.

Si A est irréductible, montrer que Σ_A est ou bien fini et réduit à une orbite périodique, ou bien un espace de Cantor.

Exercice E.31. 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. Rappelons (voir par exemple [BenP, §9]) le résultat classique suivant.

Théorème 6.22. (Théorème de Perron-Frobenius) 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} .
- b) Si $\lambda_{\max} = \lambda_{\max}(B)$ est la valeur propre de B correspondant à v_{\max} , alors λ_{\max} est simple, strictement positive, et toute autre valeur propre λ de B vérifie

$$|\lambda| < \lambda_{\max} . \quad \square$$

L'entropie topologique d'un sous-décalage de type fini est donnée par le résultat suivant.

Proposition 6.23. 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 matricielle¹⁵² $\|\cdot\|$), alors

$$h_{\text{top}}(\sigma_A) = h_{\text{top}}(\sigma_{A,+}) = \ln \rho(A) .$$

Si A est apériodique, alors $h_{\text{top}}(\sigma_A) = h_{\text{top}}(\sigma_{A,+}) = \ln \lambda_{\max}(A)$.

149. Voir l'exercice E.37 pour une explication de la terminologie.

150. 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 $(A^n)_{ij} \neq 0$.

151. Par exemple, $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est irréductible, mais pas apériodique.

152. 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 $\|AB\| \leq \|A\| \|B\|$

Démonstration. La seconde assertion découle de la première et du théorème de Perron-Frobenius. 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 6.24.

(1) Pour tout α 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 $\text{tr } A^N = \sum_{a \in \mathcal{A}} (A^N)_{a,a}$. \square

Pour tout $n \in \mathbb{N}$, appelons dans cette démonstration $(2n+1)$ -cylindre toute partie de Σ_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 Σ_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 Σ_A . Le recouvrement ouvert \mathcal{U} de Σ_A par les 1-cylindres est générateur pour σ_A , car pour tout $n \in \mathbb{N}$, 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

$$\frac{1}{c} \|B\| \leq \sum_{\alpha, \beta \in \mathcal{A}} |B_{\alpha, \beta}| \leq c \|B\| . \quad (83)$$

Par le théorème 6.9, par la formule (81), par le lemme 6.24 (1) et par la formule (83), nous avons donc

$$\begin{aligned} h_{\text{top}}(\sigma_A) &= h(\sigma_A, \mathcal{U}) = \lim_{n \rightarrow \infty} \frac{1}{2n+1} \ln N\left(\bigvee_{i=-n}^n \sigma_A^{-i} \mathcal{U}\right) = \lim_{n \rightarrow \infty} \frac{1}{2n+1} \ln \sum_{a, b \in \mathcal{A}} (A^{2n+1})_{a,b} \\ &= \lim_{n \rightarrow \infty} \frac{1}{2n+1} \ln \|A^{2n+1}\| = \ln \rho(A) . \quad \square \end{aligned}$$

6.4.2 Chaînes de Markov sur un alphabet fini

Si A est une matrice d'incidence, il existe sauf cas particulier de très nombreuses mesures invariantes sur les sous-décalages de type fini (Σ_A, σ_A) et $(\Sigma_{A,+}, \sigma_{A,+})$. Le but de cette sous-partie est de donner une construction d'une famille de telles mesures, classique en théorie des probabilités, à l'aide de matrices de transition, de calculer leur entropie mesurée et d'étudier leur propriété de mélange.

Soit \mathcal{A} un ensemble fini, de cardinal au moins 2, muni de la topologie discrète. Une *matrice de transition* ¹⁵³ 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 α dans \mathcal{A} , nous avons

$$\sum_{\beta \in \mathcal{A}} \pi_{\alpha\beta} = 1.$$

La *matrice d'incidence associée* à la matrice de transition π est la matrice d'incidence $A_\pi = (A_{\alpha\beta})_{\alpha,\beta \in \mathcal{A}}$ définie par $A_{\alpha\beta}$ vaut 1 si et seulement si $\pi_{\alpha\beta}$ est non nul. Notons que π est apériodique (respectivement irréductible) si et seulement si A l'est.

Une mesure (borélienne) de probabilité ν 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 π -stationnaire* (ou *mesure stationnaire* lorsque π est sous-entendue) sur \mathcal{A} est par définition une mesure de probabilité ν sur \mathcal{A} qui est un vecteur propre de ${}^t\pi$ pour la valeur propre 1. avec ν identifié à une matrice ligne, ceci revient à demander que l'égalité matricielle $\nu\pi = \nu$ soit vérifiée, c'est-à-dire que, pour tout β dans \mathcal{A} , nous ayons

$$\sum_{\alpha \in \mathcal{A}} \nu_\alpha \pi_{\alpha\beta} = \nu_\beta.$$

Lemme 6.25. *Soit \mathcal{A} comme ci-dessus.*

- (1) *Toute matrice de transition π sur \mathcal{A} admet au moins une mesure π -stationnaire.*
- (2) *Si π est apériodique, alors cette mesure π -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 ν^π est de support égal à \mathcal{A} (tous les coefficients de ν^π sont strictement positifs). Le réel 1 est une valeur propre simple de π et les valeurs propres λ 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 π^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 π -stationnaire.

¹⁵³. Certains ouvrages disent aussi *matrice stochastique*, et dans [BenP] c'est la matrice transposée qui est appelée une matrice stochastique.

(2) La matrice ${}^t\pi$ est aussi une matrice apériodique. Puisque π^n est une matrice de transition, la norme matricielle de $({}^t\pi)^n$ 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 6.22 appliqué à ${}^t\pi$, puisque $\lambda_{\max}({}^t\pi) = \rho({}^t\pi) = 1$. \square

Par une construction classique en théorie des probabilités, la donnée d'une matrice de transition π et d'une mesure π -stationnaire ν sur \mathcal{A} définit une marche aléatoire sur le graphe \mathcal{G}_{A_π} , de probabilité de passage du sommet α au sommet β égale à $\pi_{\alpha\beta}$, et de loi initiale ν , et donc une mesure de probabilité sur l'espace de probabilité produit $(\mathcal{A}, \nu)^{\mathbb{N}}$ des trajectoires de la marche aléatoire. Mais nous allons démontrer cela de manière autonome.

Lemme 6.26. *Pour toute matrice de transition π et toute mesure π -stationnaire ν sur \mathcal{A} , il existe une et une seule mesure de probabilité \mathbb{P}_ν (respectivement $\mathbb{P}_{\nu,+}$) sur l'espace mesurable produit $\Sigma = \mathcal{A}^{\mathbb{Z}}$ (respectivement $\Sigma_+ = \mathcal{A}^{\mathbb{N}}$) invariante par le décalage σ (respectivement σ_+) telle que, pour tous les $\alpha_0, \dots, \alpha_n \in \mathcal{A}$,*

$$\mathbb{P}_\nu(\{\omega \in \Sigma : \omega_0 = \alpha_0, \dots, \omega_n = \alpha_n\}) = \nu_{\alpha_0} \prod_{i=1}^n \pi_{\omega_{i-1}, \omega_i}.$$

(respectivement

$$\mathbb{P}_{\nu,+}(\{\omega \in \Sigma_+ : \omega_0 = \alpha_0, \dots, \omega_n = \alpha_n\}) = \nu_{\alpha_0} \prod_{i=1}^n \pi_{\omega_{i-1}, \omega_i}.$$

De plus, le support de \mathbb{P}_ν (respectivement $\mathbb{P}_{\nu,+}$) est le sous-décalage de type fini Σ_{A_π} (respectivement $\Sigma_{A_\pi,+}$).

Nous appellerons \mathbb{P}_ν (respectivement $\mathbb{P}_{\nu,+}$) la *mesure de Markov* sur Σ_{A_π} (respectivement $\Sigma_{A_\pi,+}$) associée à (π, ν) . L'application σ_{A_π} (respectivement $\sigma_{A_\pi,+}$) est parfois appelée le *décalage de Markov* du système dynamique mesuré $(\Sigma_{A_\pi}, \mathbb{P}_\nu, \sigma_{A_\pi})$ (respectivement $(\Sigma_{A_\pi,+}, \mathbb{P}_{\nu,+}, \sigma_{A_\pi,+})$).

Démonstration. Nous ne considérons que le cas bilatère, le cas unilatère se traite de même.

Rappelons d'abord le résultat suivant, que nous admettrons (voir par exemple [Coh], [Neve, page 23]). Soit X un ensemble. Une *algèbre de Boole* 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 $X \in \mathcal{E}$.

Théorème 6.27. (Théorème de Carathéodory) *Soient \mathcal{E} une algèbre de Boole de parties d'un ensemble X et \mathcal{B} la σ -algèbre de parties de X engendrée par \mathcal{E} . Considérons $\mu : \mathcal{E} \rightarrow [0, +\infty]$ une application vérifiant*

- i) *l'application μ est σ -finie sur \mathcal{E} (c'est-à-dire X est une union dénombrable d'éléments de \mathcal{E} dont l'image par μ est finie),*
- ii) *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) 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 μ se prolonge de manière unique en une mesure σ -finie sur \mathcal{B} . Si $\mu(X) = 1$, alors son extension est une mesure de probabilité.

Deux mesures sur un espace mesurable (X, \mathcal{B}) , dont au moins une est σ -finie, et qui coïncident sur une sous-algèbre de Boole de \mathcal{B} qui engendre la σ -algèbre \mathcal{B} , sont égales. \square

En généralisant la terminologie introduite dans l'exemple (2) de la partie 1.1, nous appellerons *cylindre* dans $\Sigma = \mathcal{A}^{\mathbb{Z}}$ toute partie de Σ de la forme

$$[\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_k = \{\omega \in \Sigma : \forall i = m, m+1, \dots, n, \omega_i = \alpha_i\},$$

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 μ sur Σ invariante par le décalage σ , et pour tout $k \in \mathbb{Z}$, nous avons

$$\mu([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_k) = \mu(\sigma^{-k}([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_0)) = \mu([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_0).$$

L'ensemble \mathcal{E}_{cyl} des unions finies disjointes de cylindres est une algèbre de Boole engendrant la σ -algèbre produit de Σ .

Afin de construire une mesure de probabilité borélienne \mathbb{P}_ν sur Σ , nous la définissons sur tous les cylindres en posant (avec la convention usuelle qu'un produit vide est égal à 1), pour tous les $m \leq n$ dans \mathbb{Z} et $\alpha_m, \dots, \alpha_n$ dans \mathcal{A} ,

$$\mathbb{P}_\nu([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_m) = \nu_{\alpha_m} \prod_{i=m}^{n-1} \pi_{\alpha_i \alpha_{i+1}} \in [0, +\infty[.$$

Nous étendons par additivité cette application \mathbb{P}_ν sur les 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 π est égale à 1, nous avons

$$\mathbb{P}_\nu([\alpha_m, \alpha_{m+1}, \dots, \alpha_n]_m) = \sum_{\alpha \in \Lambda} \mathbb{P}_\nu([\alpha, \alpha_m, \dots, \alpha_n]_{m-1}) = \sum_{\alpha \in \Lambda} \mathbb{P}_\nu([\alpha_m, \dots, \alpha_n, \alpha]_m).$$

Il est alors immédiat de montrer que l'application $\mathbb{P}_\nu : \mathcal{E}_{\text{cyl}} \rightarrow [0, +\infty[$ vérifie les propriétés **i)**, **ii)** et **iii)** du théorème de Carathéodory. 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 ν est une mesure de probabilité sur \mathcal{A} , nous avons

$$\mathbb{P}_\nu(X) = \mathbb{P}_\nu\left(\bigsqcup_{\alpha \in \mathcal{A}} [\alpha]_0\right) = \sum_{\alpha \in \mathcal{A}} \mathbb{P}_\nu([\alpha]_0) = \sum_{\alpha \in \mathcal{A}} \nu_\alpha = 1.$$

Enfin, la condition **iii)** du théorème de Carathéodory est automatiquement satisfaite, car les unions finies de cylindres sont des compacts de Σ , donc si $(C_n)_{n \in \mathbb{N}}$ est une suite décroissante dans \mathcal{E}_{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}_ν par le décalage est satisfaite par la dernière affirmation du théorème de Carathéodory. \square

Le résultat suivant calcule l'entropie mesurée des systèmes de Markov bilatère $(\Sigma, \mathbb{P}_\nu, \sigma)$ et unilatère $(\Sigma_+, \mathbb{P}_{\nu,+}, \sigma_+)$.

Proposition 6.28. Soient π une matrice de transition et ν une mesure π -stationnaire sur \mathcal{A} . Alors

$$h_{\mathbb{P}_\nu}(\sigma) = h_{\mathbb{P}_{\nu,+}}(\sigma_+) = - \sum_{\alpha, \beta \in \mathcal{A}} \nu_\alpha \pi_{\alpha\beta} \ln(\pi_{\alpha\beta}).$$

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. Notons ξ_0 la partition mesurable $\xi_0 = \{[\alpha]_0 : \alpha \in \Lambda\}$ de $\mathcal{A}^{\mathbb{Z}}$ par les cylindres de longueur 1 partant en 0. Cette partition est génératrice pour la transformatin inversible σ , 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 σ -algèbre produit de Σ . Nous avons donc, d'après le théorème de Kolmogorov-Sinai,

$$h_{\mathbb{P}_\nu}(\sigma) = h_{\mathbb{P}_\nu}(\sigma, \xi_0) = \lim_{n \rightarrow \infty} \frac{1}{2n+1} H_{\mathbb{P}_\nu} \left(\bigvee_{i=-n}^n \sigma^{-i} \xi_0 \right).$$

Or pour tout $N \in \mathbb{N} - \{0, 1\}$, en utilisant l'égalité

$$\ln \mathbb{P}_\nu([\alpha_{-n}, \dots, \alpha_n]_{-n}) = \ln \left(\nu_{\alpha_{-n}} \prod_{k=-n}^{n-1} \pi_{\alpha_k, \alpha_{k+1}} \right) = \ln \nu_{\alpha_{-n}} + \sum_{k=-n}^{n-1} \ln \pi_{\alpha_k, \alpha_{k+1}}$$

et le fait que $\nu \pi^p = \nu$ pour tout $p \in \mathbb{N}$,¹⁵⁴ et que la somme des coefficients de chaque ligne de π^p est égale à 1,¹⁵⁵ nous avons

$$\begin{aligned} H_{\mathbb{P}_\nu} \left(\bigvee_{i=-n}^n \sigma^{-i} \xi_0 \right) &= - \sum_{\alpha_{-n}, \dots, \alpha_n \in \mathcal{A}} \mathbb{P}_\nu([\alpha_{-n}, \dots, \alpha_n]_{-n}) \ln \mathbb{P}_\nu([\alpha_{-n}, \dots, \alpha_n]_{-n}) \\ &= - \sum_{\alpha_{-n} \in \mathcal{A}} \nu_{\alpha_{-n}} \ln \nu_{\alpha_{-n}} \sum_{\alpha_{-n+1}, \dots, \alpha_n \in \mathcal{A}} \prod_{i=-n}^{n-1} \pi_{\alpha_i, \alpha_{i+1}} \\ &\quad - \sum_{k=-n}^{n-1} \sum_{\alpha_k, \alpha_{k+1} \in \mathcal{A}} \left(\sum_{\alpha_{-n}, \dots, \alpha_{k-1} \in \mathcal{A}} \nu_{\alpha_{-n}} \prod_{i=-n}^{k-1} \pi_{\alpha_i, \alpha_{i+1}} \right) \pi_{\alpha_k, \alpha_{k+1}} \\ &\quad \quad \quad \times \ln \pi_{\alpha_k, \alpha_{k+1}} \left(\sum_{\alpha_{k+2}, \dots, \alpha_{n-1} \in \mathcal{A}} \prod_{i=k+1}^{n-1} \pi_{\alpha_i, \alpha_{i+1}} \right) \\ &= - \sum_{\alpha \in \mathcal{A}} \nu_\alpha \ln \nu_\alpha - 2n \sum_{\alpha, \beta \in \mathcal{A}} \nu_\alpha \pi_{\alpha\beta} \ln(\pi_{\alpha\beta}). \end{aligned}$$

En divisant par $2n+1$ et en faisant tendre n vers $+\infty$, ceci montre le résultat. \square

Si une matrice de transition π sur l'alphabet \mathcal{A} est apériodique, donc admet une unique mesure π -stationnaire ν^π sur \mathcal{A} par le lemme 6.25 (2), nous noterons \mathbb{P}_π et $\mathbb{P}_{\pi,+}$ les mesures \mathbb{P}_{ν^π} et $\mathbb{P}_{\nu^\pi,+}$ respectivement. Le résultat suivant montre que les sous-décalages de type fini

154. 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 $k \in \{-n, \dots, n-1\}$ et $\alpha_k \in \mathcal{A}$

155. 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 $\sum_{\alpha_{k+2}, \dots, \alpha_{n-1} \in \mathcal{A}} \prod_{i=k+1}^{n-1} \pi_{\alpha_i, \alpha_{i+1}} = 1$ pour tous les $k \in \{-n, \dots, n-1\}$ et $\alpha_{k+1} \in \mathcal{A}$

bilatère $(\Sigma_{A_\pi}, \sigma_{A_\pi})$ et unilatère $(\Sigma_{A_\pi, +}, \sigma_{A_\pi, +})$, qui sont de mesure totale dans (Σ, σ) et (Σ_+, σ_+) pour les mesures \mathbb{P}_π et $\mathbb{P}_{\pi, +}$ respectivement, sont mélangeants pour ces mesures. L'hypothèse d'apériodicité est nécessaire, comme nous le voyons facilement en prenant pour π la matrice identité.

Lemme 6.29. *Soit π une matrice de transition apériodique sur \mathcal{A} . Alors le décalage de Bernoulli bilatère σ sur Σ est mélangeant pour la mesure de Markov \mathbb{P}_π , et le décalage de Bernoulli unilatère σ_+ sur Σ_+ 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 6.26. Puisque les cylindres engendrent la tribu borélienne de Σ , 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 3.4, il suffit donc de montrer que, pour tous les cylindres C_1, C_2 de Σ , 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 α, β dans \mathcal{A} , nous avons

$$\lim_{N \rightarrow \infty} (\pi^N)_{\alpha\beta} = \nu_\beta^\pi.$$

Or, par le lemme 6.25, le polynôme caractéristique $\chi(X) = \det(X \text{id} - {}^t \pi)$ 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 τ est égale à 1. Les coefficients matriciels de τ sont donc $\tau_{\alpha\beta} = \nu_\alpha^\pi$ pour tous les $\alpha, \beta \in \mathcal{A}$, et $H_0 = \{v \in \mathbb{R}^{\mathcal{A}} : \sum_{\alpha \in \mathcal{A}} v_\alpha = 0\}$. Donc $\lim_{N \rightarrow \infty} (\pi^N)_{\alpha\beta} = {}^t \tau_{\alpha\beta} = \tau_{\beta\alpha} = \nu_\beta^\pi$ pour tous les $\alpha, \beta \in \mathcal{A}$, comme voulu. \square

Remarque. Avec l'exercice E.38, ceci montre en particulier qu'un système dynamique symbolique $(\Lambda^{\mathbb{Z}}, \sigma)$ sur un alphabet fini à au moins deux éléments admet une famille non dénombrable, à plusieurs paramètres réels, de mesures de probabilité invariantes, mélangeantes et deux à deux étrangères.

6.4.3 Problématique de l'équidistribution

Soit X un espace topologique localement compact, de tribu borélienne \mathcal{B} , muni d'une mesure (borélienne positive) régulière μ . 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 μ_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, μ 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}_μ l'ensemble des boréliens $B \in \mathcal{B}$ d'adhérence compacte tels que $\mu(\partial B) = 0$.¹⁵⁶ Des résultats classiques de théorie de la mesure (voir par exemple [Coh] ou le lemme de Portmanteau de [Bil]) montrent le résultat suivant.

Proposition 6.30. *Les conditions suivantes sont équivalentes :*

- (1) *la suite de mesures $(\mu_N)_{N \in \mathbb{N}}$ converge vaguement vers la mesure μ (c'est-à-dire (voir la partie 1.2) 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}_μ qui engendre la σ -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 μ à l'échelle $(\lambda_N)_{N \in \mathbb{N}}$.

Lorsque X est compact, si F_N est fini et $\lambda_N = \text{Card } F_N$, si les mesures μ_f pour $f \in F_N$ sont des mesures de probabilité, la mesure μ , 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 μ 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, μ) 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, μ) 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 μ (c'est-à-dire que $(\mu_N)_{N \in \mathbb{N}}$ converge vaguement vers la mesure μ) est alors équivalent à dire que pour tout

¹⁵⁶. Il est important de ne pas oublier cette condition.

élément ¹⁵⁷ $B \in \mathcal{B}_\mu$ (ou, de manière équivalente, pour tout élément B d'une partie \mathcal{B}' donnée de \mathcal{B}_μ qui engendre la σ -algèbre \mathcal{B}), la proportion de points de F_N dans B est asymptotiquement égale à la mesure de B pour μ :

$$\lim_{N \rightarrow +\infty} \frac{\text{Card}(B \cap F_N)}{\text{Card } F_N} = \mu(B).$$

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_α sur \mathbb{R}/\mathbb{Z} (et la version adaptée du théorème ergodique de Birkhoff, voir la proposition 1.8), 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 3). ¹⁵⁸

(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.

157. attention à ne pas oublier de vérifier l'hypothèse que $\mu(\partial B) = 0$

158. 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 3.1 (1).

6.4.4 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 6.22 de Perron-Frobenius, notons $\lambda = \lambda_{\max}(A)$, $v = (v_\alpha)_{\alpha \in \mathcal{A}}$ un vecteur propre de tA pour la valeur propre λ à coefficients strictement positifs, et $f = (f_\alpha)_{\alpha \in \mathcal{A}}$ un vecteur propre de A pour la valeur propre λ à coefficients strictement positifs, normalisés¹⁵⁹ de sorte que

$$\sum_{\alpha \in \mathcal{A}} v_\alpha f_\alpha = 1. \quad (84)$$

Nous avons donc, pour tout α dans \mathcal{A} ,

$$v_\alpha > 0, \quad f_\alpha > 0, \quad \sum_{\beta \in \mathcal{A}} v_\beta A_{\beta\alpha} = \lambda v_\alpha, \quad \text{et} \quad \sum_{\beta \in \mathcal{A}} A_{\alpha\beta} f_\beta = \lambda f_\alpha. \quad (85)$$

Notons alors $\pi^\bullet = (\pi_{\alpha\beta}^\bullet)_{\alpha,\beta \in \mathcal{A}}$ la matrice de coefficients

$$\pi_{\alpha\beta}^\bullet = \frac{A_{\alpha\beta} f_\beta}{\lambda f_\alpha}. \quad (86)$$

Il est immédiat de vérifier par cette formule (86) que π^\bullet est une matrice de transition apériodique, dont la matrice d'incidence associée est égale à A . La mesure de Markov associée

$$\mu = \mathbb{P}_{\pi^\bullet} \quad (87)$$

sur Σ_A , respectivement

$$\mu_+ = \mathbb{P}_{\pi^\bullet,+} \quad (88)$$

sur $\Sigma_{A,+}$, s'appelle la *mesure de Parry* du sous-décalage de type fini bilatère (Σ_A, σ_A) , respectivement unilatère $(\Sigma_{A,+}, \sigma_{A,+})$.

Exemple. Lorsque A est une matrice carrée de coefficients tous égaux à 1, alors λ vaut $q = \text{Card } \mathcal{A}$, $v_\alpha = f_\alpha = \frac{1}{\sqrt{q}}$ pour tout α dans \mathcal{A} , et π^\bullet est la matrice carrée de coefficients tous égaux à $\frac{1}{q}$. Le sous-décalage de type fini (Σ_A, σ_A) est alors égal au système de Bernoulli (Σ, σ) et la mesure de Parry de $(\Sigma_A, \sigma_A) = (\Sigma, \sigma)$ est alors la mesure produit $\nu_0^{\mathbb{Z}}$, pour ν_0 la mesure d'équiprobabilité sur \mathcal{A} .

Remarque 6.31. L'unique (par le lemme 6.25 (2)) mesure π^\bullet -stationnaire est

$$\nu^{\pi^\bullet} = (v_\alpha f_\alpha)_{\alpha \in \mathcal{A}},$$

car ν^{π^\bullet} est un vecteur de coordonnées strictement positives dont la somme est égale à 1 par la formule (84), et $\nu^{\pi^\bullet} \pi^\bullet = \nu^{\pi^\bullet}$ par les formules (85) et (86).

¹⁵⁹. 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.

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 6.32. *Soient (Σ_A, σ_A) un sous-décalage de type fini bilatère, de matrice d'incidence A apériodique, et μ la mesure de Parry sur Σ_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 σ_A de période divisant N , et $\mu_N = \frac{1}{\text{Card } P_N} \sum_{\omega \in P_N} \delta_\omega$, où δ_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 σ_A -invariante μ' sur Σ_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 Σ_A vers la mesure d'entropie maximale : la suite des mesures de probabilité $(\mu_N)_{N \in \mathbb{N}}$ converge vaguement vers μ .

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 μ_+ est sa mesure de Parry, alors

- μ_+ 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 σ 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 ν_0 l'équiprobabilité sur \mathcal{A} (voir aussi l'exercice E.36).

Démonstration. a) La proposition 6.28 appliquée à $\pi = \pi^\bullet$ et à $\nu = \nu^{\pi^\bullet}$, la remarque 6.31 et la formule (86) 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 (85) et (84), 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 6.23, le résultat en découle. Donc la mesure de Parry est une mesure d'entropie maximale du système dynamique topologique (Σ_A, σ_A) .

b) Soit μ' une autre mesure de probabilité borélienne σ_A -invariante d'entropie maximale sur Σ_A , donc telle que $h_{\mu'}(\sigma_A) = h_{\text{top}}(\sigma_A) = \ln \lambda$.

Si μ' est absolument continue par rapport à la mesure de Parry μ , alors il existe, d'après le théorème de Radon-Nikodym, une fonction ψ positive ou nulle dans $\mathbb{L}^1(\Sigma_A, \mu)$ telle que $\mu' = \psi\mu$. Cette fonction est σ_A -invariante μ -presque partout. Comme μ est ergodique (car mélangeante, voir le lemme 6.29), et puisque μ et μ' 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 μ' ne soit pas absolument continue par rapport à μ . Il existe donc un borélien B_* de Σ_A tel que $\mu'(B_*) > 0$ et $\mu(B_*) = 0$. Rappelons que l'ensemble \mathcal{C} des unions finies disjointes de cylindres est une algèbre de Boole engendrant la σ -algèbre produit de Σ .

Nous allons utiliser le lemme suivant, où nous notons $B\Delta C = (B \cup C) - (B \cap C)$ la *différence symétrique* de deux parties B et C .

Lemme 6.33. *Soit $(X'', \mathcal{B}'', \mu'')$ un espace de probabilité et \mathcal{C} une sous-algèbre de Boole de \mathcal{B}'' qui engendre la σ -algèbre \mathcal{B}'' . Alors, pour tous les 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 σ -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 de $(B_n)_{n \in \mathbb{N}}$, alors $\lim_{n \rightarrow +\infty} \mu(B\Delta B_n) = 0$. □

Ce lemme 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 Σ_A telles que $\lim_{p \rightarrow \infty} \mu'(B_p) = \mu'(B_*) \neq 0$ et $\lim_{p \rightarrow \infty} \mu(B_p) = 0$.

Notons ξ_0 la partition mesurable de Σ_A définie par

$$\xi_0 = \{\Sigma_A \cap [\alpha]_0 : \alpha \in \mathcal{A}, \Sigma_A \cap [\alpha]_0 \neq \emptyset\}$$

et ξ_n la partition $\xi_n = \bigvee_{i=-n}^n \sigma_A^{-i} \xi_0$. Comme ξ_0 est une partition génératrice pour la transformation inversible σ_A , d'après le théorème de Kolmogorov-Sinai (proposition 6.12 (6)) et la proposition 6.12 (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 6.34. *Soient x_1, \dots, x_k des réels strictement positifs tels que $\sum_{i=1}^k x_i = a$. Alors $-\sum_{i=1}^k x_i \ln x_i \leq a \ln \frac{k}{a}$.*

Démonstration. Par la concavité de la fonction $\varphi : t \mapsto -t \ln t$ sur $[0, 1]$, nous avons

$$\frac{1}{k} \sum_{i=1}^k -x_i \ln x_i \leq -\left(\frac{1}{k} \sum_{i=1}^k x_i\right) \ln \left(\frac{1}{k} \sum_{i=1}^k x_i\right) = \frac{a}{k} \ln \frac{k}{a}. \quad \square$$

Lemme 6.35. *Il existe $c > 0$ tel que, pour tout cylindre $C = [\alpha_m, \dots, \alpha_n]_m$ non vide de Σ_A , nous ayons $\mu(C) \geq c \lambda^{-(n-m)}$.*

Démonstration. Par la définition (87) de la mesure de Parry et le lemme 6.26 appliqué avec $\pi = \pi^\bullet$ et $\nu = \nu^{p_i^\bullet}$, par la remarque 6.31 et la formule (86), 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 $c = \min_{\alpha, \beta \in \mathcal{A}} v_\alpha f_\beta > 0$. \square

Fin de la démonstration de l'assertion b) du théorème 6.32. Fixons p et choisissons n de sorte que B_p soit une réunion finie de parties de la partition ξ_n . Nous avons alors, en utilisant les deux lemmes précédents,

$$\begin{aligned} 0 &\leq H_{\mu'}(\xi_n) - (2n+1) \ln \lambda \\ &= - \sum_{C \in \xi_n, C \subset B_p} \mu'(C) \ln \mu'(C) - \sum_{C \in \xi_n, C \subset {}^c B_p} \mu'(C) \ln \mu'(C) - (2n+1) \ln \lambda \\ &\leq \mu'(B_p) \ln \frac{\text{Card} \{C \in \xi_n : C \subset B_p\}}{\mu'(B_p)} + \mu'({}^c B_p) \ln \frac{\text{Card} \{C \in \xi_n : C \subset {}^c B_p\}}{\mu'({}^c B_p)} - (2n+1) \ln \lambda \\ &\leq \mu'(B_p) \ln \frac{\mu(B_p) \lambda^{2n}}{c \mu'(B_p)} + \mu'({}^c B_p) \ln \frac{\mu({}^c B_p) \lambda^{2n}}{c \mu'({}^c B_p)} - (2n+1) \ln \lambda \\ &= \mu'(B_p) \ln \frac{\mu(B_p)}{\mu'(B_p)} + \mu'({}^c B_p) \ln \frac{\mu({}^c B_p)}{\mu'({}^c B_p)} - \ln c - \ln \lambda. \end{aligned}$$

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 Σ_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 6.24, 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 (84). Notons¹⁶⁰ τ 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 λ sont de module strictement inférieur à λ , et comme λ est de multiplicité 1, la restriction de A à H est donc de rayon spectral strictement inférieur à λ . 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 (84), 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 (84) ainsi que $\lim_{N \rightarrow \infty} \frac{1}{\lambda^N} (A^N)_{\alpha, \beta} = f_\alpha v_\beta$. D'où par la formule (86), par une simplification télescopique du troisième terme et par le lemme 6.26, nous avons

$$\lim_{N \rightarrow \infty} \mu_N(C) = \left(\prod_{i=m}^{n-1} A_{\alpha_i \alpha_{i+1}} \right) \frac{v_{\alpha_m} f_{\alpha_n}}{\lambda^{n-m}} = f_{\alpha_m} v_{\alpha_m} \prod_{i=m}^{n-1} \pi_{\alpha_i \alpha_{i+1}}^\bullet = \mu(C).$$

Ceci démontre l'assertion c). □

6.5 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 ϕ_M , déjà moultment 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 6.4) du système dynamique (\mathbb{T}^N, ϕ_M) par un sous-décalage de type fini (Σ_A, σ_A) , défini à l'aide de "partitions de Markov" construites à partir des variétés stables et instables de (\mathbb{T}^N, ϕ_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 6.4 pour montrer que la mesure de Haar sur \mathbb{T}^N est l'unique mesure d'entropie maximale pour ϕ_M , pour calculer l'entropie topologique de (\mathbb{T}^N, ϕ_M) , et pour montrer l'équidistribution des points périodiques de ϕ_M vers la mesure de Haar sur \mathbb{T}^N .

160. 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

6.5.1 Partition de Markov du tore

Dans toute la suite de cette partie, nous reprenons les notations de la partie 5.3 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 ± 1 (pour que ϕ_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 avons $\|x_s + x_u\| = \max\{\|x_s\|, \|x_u\|\}$)¹⁶¹. 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

$$\text{ch}(M) = \max\{\|M|_{E^s}\|, \|M^{-1}|_{E^u}\|\} < 1 ;$$

- $\rho : \mathbb{R}^N \rightarrow \mathbb{T}^N$ la projection canonique ;
- $\phi = \phi_M$ l'automorphisme linéaire hyperbolique du tore \mathbb{T}^N défini par $\phi_M \circ \rho = \rho \circ M$; et
- d la distance quotient sur \mathbb{T}^N

$$d(x, y) = \inf\{\|v - w\| : \rho(v) = x, \rho(w) = y\} .$$

Fixons $\epsilon_0 > 0$ tel que l'application ρ 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

$$\epsilon_1 = 2\epsilon_0 \max\{1, \|M\|, \|M^{-1}\|\} . \quad (89)$$

Elle est alors isométrique sur toute boule fermée de rayon ϵ_1 .

Pour $\epsilon < \epsilon_0$ et x dans \mathbb{T}^N , notons, comme cas particulier de la situation générale décrite dans la partie 5.8,

$$W_\epsilon^s(x) = \{y \in \mathbb{T}^N : \forall n \geq 0, d(\phi^n(x), \phi^n(y)) \leq \epsilon\}$$

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.

Lemme 6.36. *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 $W_\epsilon^u(\phi(x)) \subset \phi(W_\epsilon^u(x))$.*

161. Attention, il ne s'agit pas d'une norme euclidienne.

(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 ϕ_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 5.8), 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 ϕ_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.26).

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

$$W_\epsilon^s(x) = \mathcal{P}(\overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s)) \quad \text{et} \quad W_\epsilon^u(x) = \mathcal{P}(\overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^u)) , \quad (90)$$

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 $v \in \overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s)$. 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

$$\mathcal{P}(\overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s)) \subset \{y \in \mathbb{T}^N : \forall n \in \mathbb{N}, d(\phi^n(x), \phi^n(y)) \leq \epsilon \operatorname{ch}(M)^n\} \subset W_\epsilon^s(x) . \quad (91)$$

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 (89) de ϵ_1 , l'application ρ 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 (91) est en fait une chaîne d'égalités. L'assertion (2) en découle, ainsi que l'égalité de gauche dans la formule (90).

(3) L'assertion (3), ainsi que l'égalité de droite dans la formule (90), découlent de la démonstration précédente, en remplaçant ϕ par $\phi^{-1} = \phi_{M^{-1}}$, et en remarquant que $\text{ch}(M) = \text{ch}(M^{-1})$ et que la définition (89) de ϵ_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

$$\max \{ \|\tilde{x} - \tilde{z}\|, \|\tilde{z} - \tilde{y}\| \} = \|(\tilde{x} - \tilde{z}) + (\tilde{z} - \tilde{y})\| = \|\tilde{x} - \tilde{y}\| \leq \epsilon$$

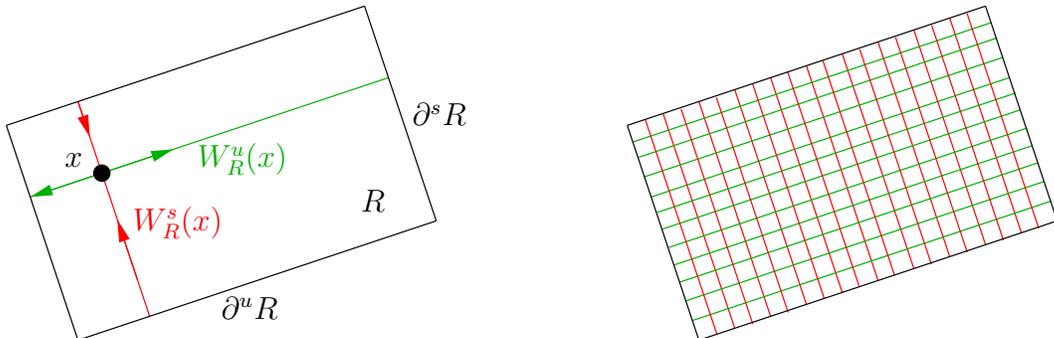
Par les formules (90), nous avons donc $z = \rho(\tilde{z}) \in W_\epsilon^s(x) \cap W_\epsilon^u(y)$. Si z' est un autre point de $W_\epsilon^s(x) \cap W_\epsilon^u(y)$, alors z' admet un unique relevé \tilde{z}' appartenant à $\overline{B}(\tilde{x}, \epsilon) \cap (\tilde{x} + E^s)$ et il existe un 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 ρ sur le même point y . Puisque ρ est injective sur la boule $\overline{B}(\tilde{x}, \epsilon)$ par la définition (89) de ϵ_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 ϵ_0 telle que si $x, y \in R$, alors $[x, y] \in R$.

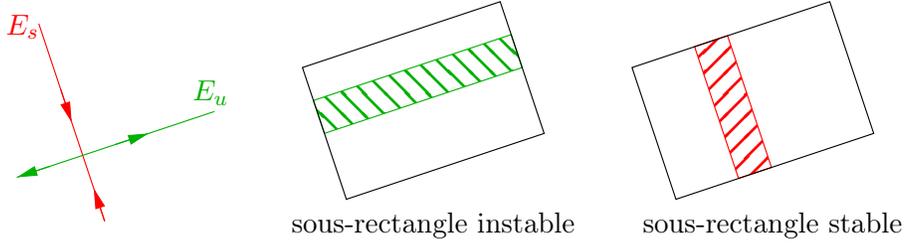
Autrement dit, dans une identification locale de \mathbb{T}^N avec $E = E^s \times E^u$, nous avons $R = R^s \times R^u$ avec $\text{diam}(R^s) \leq \epsilon_0$ et $\text{diam}(R^u) \leq \epsilon_0$. Les parties R^s et R^u ne sont pas forcément des boules de E^s et E^u . La frontière de R s'écrit alors $\partial R = \partial^s R \cup \partial^u R$ 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

$$W_R^s(x) = W_{\epsilon_0}^s(x) \cap R \quad \text{et} \quad W_R^u(x) = W_{\epsilon_0}^u(x) \cap R,$$

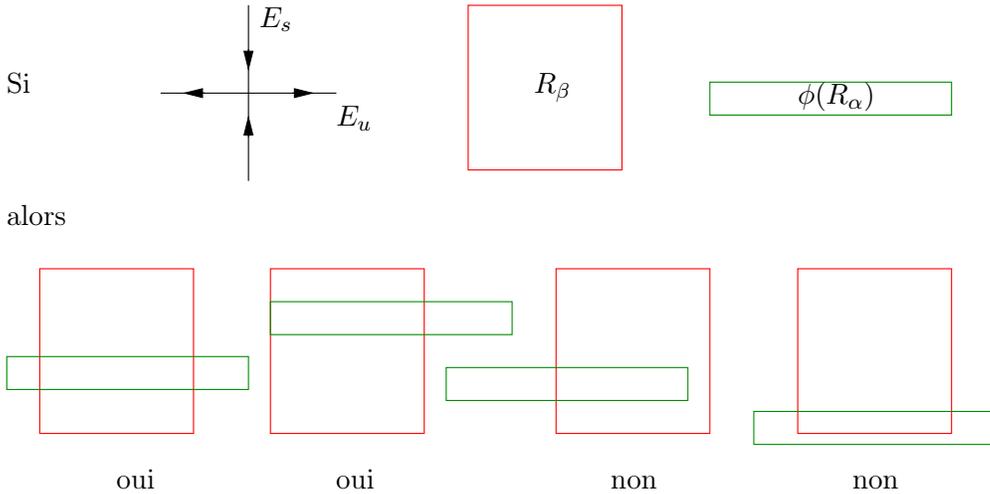
que nous appellerons les feuilles stables et instables dans R de x . Une partie S de R est appelée un *sous-rectangle stable* (resp. *instable*) si pour tout x dans S , la partie $W_R^s(x)$ (resp. $W_R^u(x)$) est contenue dans S .



Une *partition de Markov*¹⁶² \mathcal{R} de \mathbb{T}^N pour ϕ 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 ϕ , tel que, pour tous les α, β 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 6.37. *Pour tout $\epsilon > 0$ et tout automorphisme linéaire hyperbolique ϕ_M de \mathbb{T}^N , il existe une partition de Markov \mathcal{R} pour ϕ_M par des rectangles de diamètre au plus ϵ . \square*

162. 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.

6.5.2 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 6.38. *Soient $\mathcal{R} = (R_\alpha)_{\alpha \in \mathcal{A}}$ une partition de Markov pour $\phi = \phi_M$ par des rectangles de diamètre au plus ϵ avec $\epsilon > 0$ assez petit, A sa matrice d'incidence et (Σ_A, σ_A) le sous-décalage de type fini associé.*

- Pour tout ω dans Σ_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 σ_A et ϕ .
- Pour toute mesure de probabilité σ_A -invariante et ergodique μ , de support Σ_A , nous avons

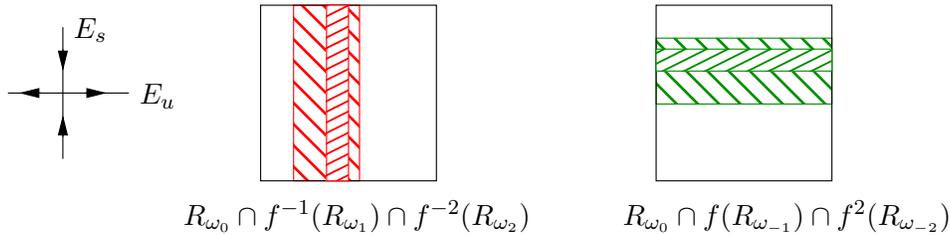
$$\mu(\{\omega \in \Sigma_A : \text{Card } h^{-1}(h(\omega)) > 1\}) = 0.$$

Ce théorème (avec l'existence de partitions de Markov vue en théorème 6.37) 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 Σ_A . Par la définition de Σ_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_{ω_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 ϵ près (puisque les diamètres des rectangles sont au plus ϵ), et sont donc égaux (par le lemme de pistage 5.20) si ϵ 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 Σ_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 ϕ , 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 μ' est invariante par ϕ , 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 μ' 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 3.5 et l'exercice E.13), le sous-décalage σ_A est aussi topologiquement mélangeant et la matrice A est irréductible (voir l'exercice E.37). 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 σ_A et ergodique \mathbb{P}_Π de support Σ_A . Par l'exercice E.38, 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 ϕ_M -invariantes, ergodiques et deux à deux étrangères.

6.5.3 Mesure d'entropie maximale des automorphismes hyperboliques du tore

Proposition 6.39. Soient ϕ_M un automorphisme linéaire hyperbolique du tore \mathbb{T}^N et μ 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, μ est l'unique mesure d'entropie maximale de ϕ_M sur \mathbb{T}^N .

Remarque. Nous avons $\ln \lambda_+ = \sum \ln |\lambda_i|$, où les λ_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 ϕ_M . Choisissons $\xi_0 = \{R'_\alpha : \alpha \in \mathcal{A}\}$ une (vraie) partition mesurable de \mathbb{T}^N telle que, pour tout α , nous ayons $\overset{\circ}{R}_\alpha \subset R'_\alpha \subset R_\alpha$. La proposition résulte alors des deux lemmes suivants.

Lemme 6.40. *Il existe $c_1, c_2 > 0$ tels que, pour tout $n \geq 0$ et tout évènement C non μ -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 μ par la projection canonique \mathcal{p} . 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_α est propre, les intérieurs de R_α^s et R_α^u sont non vides, donc $m_s(R_\alpha^s) > 0$ et $m_u(R_\alpha^u) > 0$. De plus, pour tout w dans Σ_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 6.41. *Soient X un espace métrique compact, ϕ un homéomorphisme de X , μ une mesure de probabilité borélienne ϕ -invariante et ergodique sur X , et ξ_0 une partition mesurable de X génératrice pour ϕ . Supposons qu'il existe $c_1, c_2, \lambda_+ > 0$ tels que, pour tout évènement C non μ -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 $h_{\text{top}}(\phi) = h_\mu(\phi) = \ln \lambda_+$. En outre, μ est l'unique mesure d'entropie maximale de ϕ sur X .

Démonstration. Par la définition de l'entropie d'une partition mesurable β 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 6.13), il reste à montrer que si une mesure de probabilité borélienne ϕ -invariante μ' 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 6.32 b). \square

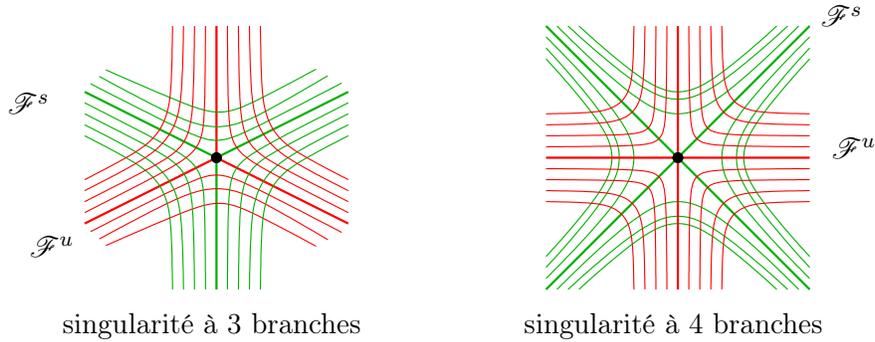
Reprenons la démonstration de la proposition 6.39. Nous avons vu que ϕ_M est mélangeante, donc ergodique, pour la mesure de Haar μ (voir le théorème 3.5). La partition ξ_0 construite en début de démonstration est génératrice pour ϕ_M (qui est inversible), car par

la démonstration de l'assertion **b)** du théorème 6.38, les événements 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 σ -algèbre des boréliens. Le lemme 6.41 précédant permet alors de conclure la démonstration de la proposition 6.39. \square

Ainsi, si $N = 2$, et si λ, λ' 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 ϕ 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 ϕ) et (\mathcal{F}^s, μ^s) et (\mathcal{F}^u, μ^u) deux feuilletages à singularités isolées de type selle à au moins trois branches, invariants par ϕ 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 ϕ_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 ϕ_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.34, 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 5.9).

Théorème 6.42. (Théorème de Bowen) *Soit Λ une pièce basique d'un C^1 -difféomorphisme ϕ Axiome A d'une variété lisse M . Alors $\phi|_{\Lambda}$ admet une et une seule mesure d'entropie maximale μ_{Λ} . Celle-ci est ergodique. Les points périodiques de $\phi|_{\Lambda}$ s'équidistribuent*

vers μ_Λ : 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 δ_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

6.6 Exercices

Exercice E.32. (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 ϕ 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.33. Soient X et Y des espaces métriques compacts. Considérons des applications continues $f : X \rightarrow X$, $g : Y \rightarrow Y$ et $h : X \rightarrow Y$ avec h surjective, telles que $h \circ f = g \circ h$.¹⁶³ Montrer que

$$h_{\text{top}}(g) \leq h_{\text{top}}(f) .$$

Exercice E.34. Soient X un espace métrique compact et $\phi : X \rightarrow X$ une application continue. Montrer que si ϕ admet une unique mesure d'entropie maximale μ , alors μ est un point extrémal du convexe $\text{Prob}_\phi(X)$, et donc que ϕ est ergodique pour μ .

Exercice E.35. 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 ϕ_N .

- (1) Soit μ une mesure de probabilité borélienne invariante par ϕ_N sur le cercle X . Montrer que $h_\mu(\phi_N) \leq \ln N$.¹⁶⁴
- (2) Montrer qu'il y a égalité si et seulement si μ est la mesure de Lebesgue λ sur X .
- (3) Calculer l'entropie topologique de ϕ_N .

Exercice E.36. 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 ν 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 σ_+ .

- (1) Soit μ une mesure de probabilité borélienne invariante par σ_+ sur Σ_+ . Montrer que $h_\mu(\sigma_+) \leq \ln N$.
- (2) Montrer qu'il y a égalité si et seulement si μ est la mesure produit $\nu^{\mathbb{N}}$ sur X .

163. L'application h est donc une semi-conjugaison entre les systèmes dynamiques topologiques (X, f) et (Y, g) .

164. Plus généralement, montrer que si $(X, \mathcal{B}, \mu, \phi)$ est un système dynamique mesuré, où X est un espace métrique compact et μ une mesure de probabilité, et si X admet une partition mesurable génératrice pour ϕ de cardinal N , alors $h_\mu(\phi) \leq \ln N$.

- (3) Calculer l'entropie topologique de σ_+ .
- (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.37. Soit $A = (A_{\alpha,\beta})_{\alpha,\beta \in \mathcal{A}}$ une matrice carrée de coefficients égaux à 0 ou 1, qui est irréductible (voir l'exercice E.30). Nous appellerons *période* d de A le plus grand diviseur commun de l'ensemble $\{n > 0 : \exists \alpha \in \Lambda, (A^n)_{\alpha\alpha} \neq 0\}$. Montrer que A est apériodique si et seulement si A est irréductible de période 1.

Exercice E.38. Soient π et π' deux matrices stochastiques apériodiques distinctes, indexées par un même alphabet fini \mathcal{A} . Montrer que les mesures de Markov \mathbb{P}_π et $\mathbb{P}_{\pi'}$ sur $\mathcal{A}^{\mathbb{Z}}$ sont étrangères.

Exercice E.39. Soient (Σ_A, σ_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 σ_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.40. Soient $M \in \text{SL}_2(\mathbb{Z})$ une matrice hyperbolique, de valeurs propres λ_1, λ_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 ϕ_M de période divisant n .

- (1) Montrer qu'un point de $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ est périodique pour ϕ_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

$$\text{Card } P_n \sim e^{n h_{\text{top}}(\phi_M)} .$$

6.7 Indications pour la résolution des exercices

Exercice E.28 Montrons que, pour tout recouvrement ouvert \mathcal{U} de X , nous avons

$$h(\phi^{-1}, \mathcal{U}) = h(\phi, \mathcal{U}).$$

En prenant la borne supérieure sur les recouvrements ouverts \mathcal{U} , cela montrera le résultat. Il découle du lemme 6.1 i) que pour tout $m \in \mathbb{Z}$, nous avons $N(\phi^m \mathcal{U}) = N(\mathcal{U})$. Donc

$$h(\phi^{-1}, \mathcal{U}) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln N\left(\bigvee_{i=0}^{n-1} \phi^i \mathcal{U}\right) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln N\left(\phi^{n-1} \bigvee_{i=0}^{n-1} \phi^{-i} \mathcal{U}\right) = h(\phi, \mathcal{U}).$$

Exercice E.30 Il est immédiat que si A est irréductible, alors A n'a pas de ligne ou de colonne nulle, sinon A^n aurait la même ligne ou colonne nulle pour tout $n \in \mathbb{N} - \{0\}$.

L'équivalence entre (1) et (2) découle du fait que pour tous les $\alpha, \beta \in \mathcal{A}$ et $k \in \mathbb{N} - \{0\}$, il existe un chemin orienté de k arêtes entre α et β si et seulement si $(A^k)_{\alpha\beta} \neq 0$.

L'équivalence entre (3) et (4), ainsi que l'équivalence entre (5) et (6), ont été vues dans l'exercice E.7.

Appelons *suite admissible* toute suite finie $(\alpha_0, \dots, \alpha_n)$ dans \mathcal{A} telle que $A_{\alpha_i \alpha_{i+1}} = 1$ pour $i = 0, \dots, n-1$. Toute sous-suite consécutive dans un élément de Σ_A et de $\Sigma_{A,+}$ est une suite admissible, par définition de ces espaces. Il existe un ensemble dénombrable de suites admissibles, que nous notons w_0, w_1, \dots . Si A est irréductible, il existe, pour tout $i \in \mathbb{N}$, une suite admissible commençant par la dernière lettre de w_i et finissant par la première lettre de w_{i+1} . Il est donc possible de former un élément ω_+ de $\Sigma_{A,+}$ qui contient tous les mots w_i , que nous pouvons compléter en un élément ω de Σ_A . Ceci montre que l'orbite positive de ω_+ par $\sigma_{A,+}$ est dense dans $\Sigma_{A,+}$, et que l'orbite positive de ω par σ_A est dense dans Σ_A . Donc (1) implique (4) et (6).

Supposons que l'assertion (3) soit vérifiée. Soient $\alpha, \beta \in \mathcal{A}$. Les cylindres $[\alpha]_0$ et $[\beta]_0$ sont des ouverts non vides, car A n'a pas de ligne ni de colonne nulle, donc pour tout $\gamma \in \mathcal{A}$, il existe $\omega \in \Sigma_A$ tel que $\omega_0 = \gamma$. Puisque (Σ_A, σ_A) est positivement transitif, il existe donc $n \in \mathbb{N} - \{0\}$ tel que $\sigma_A^{-n}([\beta]_0) \cap [\alpha]_0$ soit non vide. Si $(x_n)_{n \in \mathbb{Z}} \in \Sigma_A$ appartient à cette intersection, alors $x_0 = \alpha$ et $x_n = \beta$ et $A_{x_i, x_{i+1}} = 1$ pour $i = 0, \dots, n-1$, donc il existe un chemin orienté de α à β . Ceci montre l'assertion (2). La même démonstration en remplaçant \mathbb{Z} par \mathbb{N} montre que (5) implique (2). Nous avons donc montré l'équivalence entre les assertions (1) à (6).

Rappelons qu'un *espace de Cantor* est un espace topologique métrisable compact, totalement discontinu, sans point isolé (et que deux tels espaces sont homéomorphes).

Pour toute matrice d'incidence A , les espaces Σ_A et $\Sigma_{A,+}$ sont métrisables compacts et totalement discontinus, car fermés dans les espaces métrisables compacts et totalement discontinus Σ et Σ_+ . Si A est irréductible, ou bien il existe un sommet $\alpha \in \mathcal{A}$ dont partent au moins deux arêtes distinctes, ou bien il existe un unique cycle orienté de longueur $n \geq \text{Card } \mathcal{A}$ passant par tous les sommets du graphe orienté \mathcal{G}_A et ne suivant pas deux fois la même arête. Dans le second cas, Σ_A et $\Sigma_{A,+}$ sont réduits à une orbite périodique de période n du décalage σ_A et $\sigma_{A,+}$ respectivement. Dans le premier cas, pour tout $\omega \in \Sigma_A$ (respectivement $\omega \in \Sigma_{A,+}$), pour tout $n \in \mathbb{N}$, soit $\underline{x} = (x_0, \dots, x_k)$ une suite admissible telle que $x_0 = \omega_n$ et $x_k = \alpha$. Alors l'un des (au moins) deux mots ω' qui coïncide avec ω jusqu'au temps n , puis suit la suite admissible \underline{x} , puis vaut au temps $n+k+1$ le sommet

terminal de l'une des (au moins) deux arêtes distinctes partant de α , est différent de ω et à distance au plus e^{-n} de ω . Donc les espaces Σ_A et $\Sigma_{A,+}$ n'ont pas de point isolé, et sont des espaces de Cantor.

Exercice E.31 Si $x = (x_n)_{n \in \mathbb{N}}$ et $y = (y_n)_{n \in \mathbb{N}}$ sont deux éléments distincts de $\Sigma_{A,+}$, alors il existe $n \in \mathbb{N}$ tel que $x_n \neq y_n$. Par la définition de la distance (voir les formules (4)), nous avons $d(\sigma_{A,+}^n(x), \sigma_{A,+}^n(y)) = 1$. Donc $(\Sigma_{A,+}, \sigma_{A,+})$ est expansif, et admet 1 comme constante d'expansivité (comme défini avant la proposition 5.19). La dernière affirmation découle du corollaire 6.21.

La démonstration est la même pour (Σ_A, σ_A) en remplaçant \mathbb{N} par \mathbb{Z} .

Exercice E.32 (1) Puisque ϕ est 1-lipschitzienne, les distances d_n^ϕ et d coïncident pour tout $n \in \mathbb{N}$, donc la suite $(N_{d_n^\phi}(\epsilon))_{n \in \mathbb{N}}$ est constante. Il suffit alors d'appliquer la proposition 6.5.

Exercice E.33 La moitié de la démonstration de la proposition 6.4 convient. En effet, pour tout recouvrement ouvert \mathcal{V} de Y , le recouvrement ouvert image réciproque $\mathfrak{h}^{-1}\mathcal{V} = \{\mathfrak{h}^{-1}(V) : V \in \mathcal{V}\}$ de X vérifie $N(\mathfrak{h}^{-1}\mathcal{V}) = N(\mathcal{V})$ par le lemme 6.1 i) et la surjectivité de \mathfrak{h} . Donc pour tout recouvrement ouvert \mathcal{U} de Y , nous avons

$$\begin{aligned} h(g, \mathcal{U}) &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln N\left(\bigvee_{i=0}^{n-1} g^{-i}\mathcal{U}\right) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln N\left(\mathfrak{h}^{-1}\bigvee_{i=0}^{n-1} g^{-i}\mathcal{U}\right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \ln N\left(\bigvee_{i=0}^{n-1} f^{-i}\mathfrak{h}^{-1}\mathcal{U}\right) = h(f, \mathfrak{h}^{-1}\mathcal{U}) \leq h_{\text{top}}(f). \end{aligned}$$

En prenant la borne supérieure sur les recouvrements ouverts \mathcal{U} de Y , nous avons donc $h_{\text{top}}(g) \leq h_{\text{top}}(f)$.

Exercice E.34 Soit μ une mesure d'entropie maximale pour ϕ . Supposons qu'il existe $t \in [0, 1]$ et $\nu_1, \nu_2 \in \text{Prob}_\phi(X)$ telle que $\mu = t\nu_1 + (1-t)\nu_2$. Par la dépendance affine de l'entropie en la mesure (voir la proposition 6.12 (5)), nous avons

$$h_{\text{top}}(\phi) = h_\mu(\phi) = t h_{\nu_1}(\phi) + (1-t) h_{\nu_2}(\phi) \leq t h_{\text{top}}(\phi) + (1-t) h_{\text{top}}(\phi) = h_{\text{top}}(\phi).$$

Donc si $t \neq 0, 1$, alors ν_1 et ν_2 sont des mesures d'entropie maximale, et elles sont égales à μ si μ est unique. Par conséquent, si μ est unique, alors μ est un point extrémal de $\text{Prob}_\phi(X)$, donc est ergodique par la proposition 1.6.

Exercice E.35 (1) Soit ξ la partition mesurable $\{[\frac{k}{N}, \frac{k+1}{N}[: k = 0, \dots, N-1\}$. Alors, par récurrence sur n dans $\mathbb{N} - \{0\}$, nous avons

$$\xi_n = \bigvee_{i=0}^{n-1} \phi_N^{-i}\xi = \left\{ \left[\frac{k}{N^n}, \frac{k+1}{N^n} \right[: k = 0, \dots, N^n - 1 \right\},$$

donc ξ est génératrice pour ϕ_N . Par le théorème de Kolmogorov-Sinai et la proposition 6.12 (1), pour tout n dans $\mathbb{N} - \{0\}$, nous avons

$$h_\mu(\phi_N) = h_\mu(\phi_N, \xi) = \inf_{p \geq 1} \frac{1}{p} H_\mu(\xi_p) \leq \frac{1}{n} H_\mu(\xi_n) \leq \frac{1}{n} \ln(\text{Card } \xi_n) = \ln N,$$

la dernière inégalité ayant lieu, par la proposition 6.11 (1), avec égalité si et seulement si $\mu([\frac{k}{N^n}, \frac{k+1}{N^n}[) = \frac{1}{N^n}$ pour $k = 0, \dots, N^n - 1$.

(2) Il découle de ce qui précède que si $h_\mu(\phi_N) = \ln N$, alors les mesures μ et λ coïncident sur tous les intervalles d'extrémités N -adiques (c'est-à-dire de la forme $[\frac{k}{N^n}, \frac{k+1}{N^n}[$ pour $k = 0, \dots, N^n - 1$ et n dans $\mathbb{N} - \{0\}$). Par densité, ces deux mesures sont donc égales, et la mesure de Lebesgue λ est l'unique mesure d'entropie maximale pour ϕ_N .

(3) Par le principe variationnel, l'entropie topologique de ϕ_N vaut donc

$$h_{\text{top}}(\phi_N) = \sup_{\mu \in \text{Prob}_{\phi_N}(X)} h_\mu(\phi_N) = \ln N .$$

Remarque. Il est bien entendu possible d'utiliser l'exercice E.36, l'exercice E.8 (2) et la proposition 6.20 pour résoudre les questions (1) et (2). Notons toutefois que les systèmes dynamiques topologiques $(\mathbb{R}/\mathbb{Z}, \phi_N)$ et $(\{0, \dots, N-1\}^{\mathbb{N}}, \sigma_+)$ ne sont pas (topologiquement) conjugués, car le cercle n'est pas un espace de Cantor, et donc il n'est pas possible d'utiliser directement la proposition 6.4.

Exercice E.36 (1) Soit $\xi = \{[a] = \{(x_i)_{i \in \mathbb{N}} \in \Sigma_+ : x_0 = a\} : a \in \mathcal{A}\}$ la partition mesurable de Σ_+ donnée par les cylindres de longueur 1 partant du temps 0. Alors, par récurrence sur n dans $\mathbb{N} - \{0\}$, en notant $[a_0, \dots, a_{n-1}] = \{(x_i)_{i \in \mathbb{N}} \in \Sigma_+ : x_0 = a_0, \dots, x_{n-1} = a_{n-1}\}$, nous avons

$$\xi_n = \bigvee_{i=0}^{n-1} \sigma_+^{-i} \xi = \{[a_0, \dots, a_{n-1}] : a_0, \dots, a_{n-1} \in \mathcal{A}\} ,$$

donc ξ est génératrice pour σ_+ car les cylindres engendrent la tribu borélienne de Σ_+ . Par le théorème de Kolmogorov-Sinai et la proposition 6.12 (1), pour tout n dans $\mathbb{N} - \{0\}$, nous avons

$$h_\mu(\sigma_+) = h_\mu(\sigma_+, \xi) = \inf_{p \geq 1} \frac{1}{p} H_\mu(\xi_p) \leq \frac{1}{n} H_\mu(\xi_n) \leq \frac{1}{n} \ln(\text{Card } \xi_n) = \ln N ,$$

la dernière inégalité ayant lieu, par la proposition 6.11 (1), avec égalité si et seulement si $\mu([a_0, \dots, a_{n-1}]) = \frac{1}{N^n}$ pour $a_0, \dots, a_{n-1} \in \mathcal{A}$.

(2) Il découle de ce qui précède que si $h_\mu(\sigma_+) = \ln N$, alors les mesures μ et $\nu^{\mathbb{N}}$ coïncident sur tous les cylindres. Par densité, ces deux mesures sont donc égales, et la mesure produit $\nu^{\mathbb{N}}$ est l'unique mesure d'entropie maximale pour ϕ .

(3) Par le principe variationnel, l'entropie topologique de σ_+ vaut donc

$$h_{\text{top}}(\sigma_+) = \sup_{\mu \in \text{Prob}_{\sigma_+}(X)} h_\mu(\sigma_+) = \ln N .$$

(4) La démonstration est similaire car la même partition ξ est génératrice pour le système dynamique inversible (Σ, σ) . En effet, pour tout n dans $\mathbb{N} - \{0\}$, en posant

$$[a_{1-n}, \dots, a_0, \dots, a_{n-1}] = \{(x_i)_{i \in \mathbb{Z}} \in \Sigma : x_{1-n} = a_{1-n}, \dots, x_{n-1} = a_{n-1}\} ,$$

nous avons

$$\bigvee_{i=1-n}^{n-1} \sigma^{-i} \xi = \{[a_{1-n}, \dots, a_0, \dots, a_{n-1}] : a_{1-n}, \dots, a_{n-1} \in \mathcal{A}\} ,$$

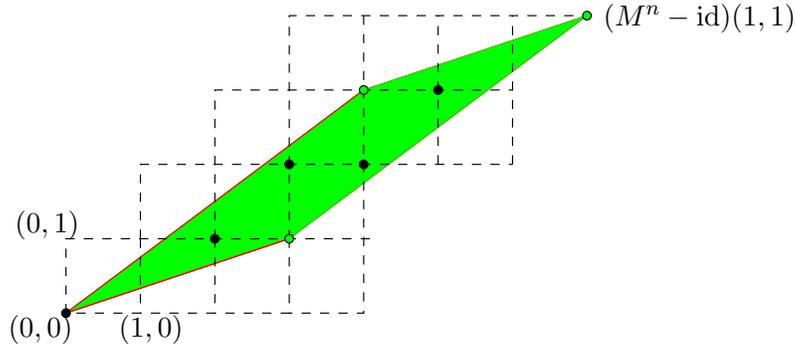
qui est de cardinal N^{2n-1} .

Exercice E.39 Soit $\lambda = \lambda_{\max}(A)$ la valeur propre de module maximal de la matrice A . Par le lemme 6.24, nous avons $\text{Card } P_N = \text{tr } A^N$. Puisque λ est strictement positive, de multiplicité 1, et que toutes les autres valeurs propres sont de module strictement inférieur à λ , nous avons $\text{tr } A^N \sim \lambda^N$. Le résultat découle alors de la proposition 6.23.

Exercice E.40 (1) Pour tout $x \in \mathbb{R}^2$, notons \bar{x} son image dans \mathbb{T}^2 . S'il existe un entier $n \in \mathbb{N} - \{0\}$ tel que $\phi_M^n(\bar{x}) = \bar{x}$, alors $M^n x - x \in \mathbb{Z}^2$. Puisque M est hyperbolique, la matrice $M^n - \text{id}$ est inversible sur \mathbb{R} et à coefficients entiers, donc $(M^n - \text{id})^{-1} \in \text{GL}_2(\mathbb{Q})$, et donc $x \in (M^n - \text{id})^{-1}(\mathbb{Z}^2) \subset \mathbb{Q}^2$. Par conséquent $\bar{x} \in \mathbb{Q}^2/\mathbb{Z}^2$.

Réciproquement, si $\bar{x} \in \mathbb{Q}^2/\mathbb{Z}^2$, alors $x \in \mathbb{Q}^2$ appartient à $(\frac{p}{q}, \frac{p'}{q}) + \mathbb{Z}^2$ avec p, p', q entiers tels que $0 \leq p, p' < q$. Par linéarité, pour tout $n \in \mathbb{N}$, nous avons $M^n x \in (\frac{pn}{q}, \frac{p'n}{q}) + \mathbb{Z}^2$ avec $0 \leq pn, p'n < q$. Par la finitude du nombre de couples d'entiers positifs au plus q , il existe donc $n' > n$ tels que $M^{n'} x - M^n x \in \mathbb{Z}^2$. Par conséquent $\phi_M^{n'-n}(\bar{x}) = \bar{x}$, et \bar{x} est un point périodique de ϕ_M .

(2) Notons Q le parallélogramme $(M^n - \text{id})([0, 1]^2)$ (avec deux côtés enlevés).



Comme vu ci-dessus, si $x \in [0, 1]^2$, nous avons l'égalité $\phi_M^n(\bar{x}) = \bar{x}$ si et seulement si $(M^n - \text{id})x \in \mathbb{Z}^2$. Donc $\text{Card } P_n = \text{Card}(Q \cap \mathbb{Z}^2)$. Ce nombre est égal à l'aire de Q , c'est-à-dire, par le théorème de changement de variable (ici linéaire), à $|\det(M^n - \text{id})|$. En diagonalisant M , et en rappelant que $\det M = \lambda_1 \lambda_2 = 1$, nous avons donc

$$\text{Card } P_n = |(\lambda_1^n - 1)(\lambda_2^n - 1)| = |\lambda_1^n + \lambda_2^n - 2| .$$

(3) Si par exemple $|\lambda_1| > 1 > |\lambda_2|$, alors, par la proposition 6.39, nous avons

$$\text{Card } P_n \sim |\lambda_1|^n = e^{n h_{\text{top}}(\phi_M)} .$$

Index

- σ -additive, 13
- algèbre
 - de Boole, 184
 - séparante, 51
- alphabet, 8, 180
- angle, 34
- application
 - à variation bornée, 38
 - analytique réelle, 44
 - contractante, 46
 - de doublement de l'angle, 24, 34
 - différentiable, 57
 - dilatante, 58
 - du chat d'Arnold, 110
 - du fer à cheval de Smale, 124
 - équicontinue, 92
 - ergodique, 15
 - expansive, 134
 - positivement, 134
 - höldérienne, 43
 - lipschitzienne, 105
 - mélangeante, 54
 - monotone, 39
 - propre, 14
 - préservant la mesure, 5
 - résolvante, 154
 - s'annulant à l'infini, 13
 - semi-continue supérieurement, 25
 - strictement dilatante, 45
 - trigonométrique, 50
 - uniformément continue, 105
 - étagée, 55
- atome, 171
- attracteur de Smale, 127
- attractif, 7
- base dénombrable d'ouverts, 17
- bassin
 - d'attraction, 127
 - de répulsion, 127
- bloc de Jordan, 104
- borné, 116
- champ
 - de cônes, 128
 - de vecteurs, 118
- chat d'Arnold, 110
- codage, 24, 179
- coefficient
 - de corrélation, 56
 - de Fourier, 51
 - matriciel, 74
- compactifié d'Alexandrov, 10
- complexifié, 155
- condition diophantienne, 45
- conjugaison, 6
 - ensembliste, 6
- conjugués, 6
 - ensemblissement, 6
- constante
 - d'expansivité, 134
 - d'hyperbolicité, 102
 - de Lipschitz, 105
- convergence faible, 77
- courbe modulaire, 90
- critère de Weyl, 189
- crochet de Bowen, 149
- cylindre, 9, 185
- décalage
 - de Bernoulli, 8
 - de Markov, 184
 - sous-décalage de type fini, 180
- degré, 34
- demi-plan supérieur, 80
- ϵ -dense, 167
- différence symétrique, 192
- différentielle, 112
- difféomorphisme
 - Anosov, 122
 - Axiome A, 122
 - de Kupka-Smale, 154
- dilatante, 45
- diophantien, 45
- distance
 - de pistage, 169
 - hyperbolique, 85
 - invariante à gauche, 69
- domaine fondamental, 85
 - strict, 85
- dual topologique, 14
- dynamique Nord-Sud, 10
- décomposition de Cartan, 77
- décroissance des corrélations, 57
 - exponentielle, 57
 - polynomiale, 57
 - rapide, 57
- dérivées partielles, 57
- endomorphisme linéaire hyperbolique, 101
- ensemble
 - α -limite, 18
 - non errant, 19
 - ω -limite, 18

- entropie, 168, 172, 173
 - mesurée, 173
 - métrique, 173
 - relative, 172
 - relativement à une partition, 173
 - relativement à un recouvrement, 168
 - topologique, 168, 170
- équidistribution, 188
- ergodique, 15
 - uniquement, 17
- errant, 19
- escalier de Lebesgue, 42
- espace
 - à l'infini, 86
 - de Cantor, 205
 - de Baire, 20
 - de Cantor, 9
 - triadique, 9
 - des phases, 5
 - vectorel topologique, 13
- évènement, 171
- expansif, 134
 - positivement, 134
- extension naturelle, 9
- facteur, 6
- fer à cheval de Smale, 124
- fibres, 114
- fibré tangent, 81
 - unitaire, 82
- fibré vectoriel, 114
 - morphisme, 114
 - normé, 115
 - rang, 114
 - restriction, 115
 - somme directe, 115
- fixateur, 74
- flot, 5
 - d'Anosov, 123
 - géodésique, 88, 89
 - horocyclique, 88, 89
 - instable, 90
 - stable, 89
- frontière, 22
- géodésique hyperbolique, 85
- groupe
 - modulaire, 83
 - topologique, 50
 - unimodulaire, 70
 - unitaire, 72
- $H_\mu(\alpha)$, 172
- $H_\mu(\alpha|\beta)$, 172
- $h(\phi, \mathcal{U})$, 168
- $h_\mu(\phi)$, 173
- $h_\mu(\phi, \alpha)$, 173
- $h_{\text{top}}(\phi)$, 168
- homéomorphisme pseudo-Anosov, 202
- horicycle, 87
- horoboule, 97
- horocycle, 87
 - instable, 87
 - stable, 87
- hyperbolique, 101, 108, 122
- instable, 101, 118, 123
- invariante, 12, 101
 - globalement, 12
 - positivement, 12
- inégalité de Hölder, 64
- joint, 166, 171
- lemme
 - de fermeture d'Anosov, 137
 - de Lebesgue, 169
 - de pistage, 135
 - linéaire, 152
- Liouville, 45
- localement maximal, 149
- longueur
 - d'arc, 85
 - hyperbolique, 85
- matrice
 - d'incidence, 180, 199
 - apériodique, 181
 - associée, 183
 - de transition, 183
 - irréductible, 180
 - période d'une, 204
 - stochastique, 183
- mélange
 - d'ordre k , 60
 - faible, 54
 - fort, 54
 - mesuré, 25, 54
 - effectif, 57
 - exponentiel, 25, 57
 - polynomial, 57
 - rapide, 57
 - multiple, 60
 - topologique, 20
- mesure
 - d'entropie maximale, 178
 - de Liouville, 83
 - de Haar, 50, 70, 72, 80

- de Lebesgue, 92
- de Liouville, 82
- de Markov, 184
- de Parry, 190
- image, 5
- mélangeante, 54
- σ -finie, 184
- signée, 13
- sphérique, 82
- stationnaire, 183
- minimal, 22
 - exceptionnel, 40
- moins fin, 167
- moins fine, 172
- mot, 8
- moyenne orbitale, 17
- $N'_d(\epsilon)$, 167
- $N_d(\epsilon)$, 167
- neutre, 123
- nombre
 - de translation, 35
 - de Liouville, 45
 - de rotation, 35
 - de type Roth, 45
 - diophantien, 45
- non errant, 19
- norme, 115, 117
 - adaptée, 102, 120
 - d'opérateur, 101, 115
 - de Sasaki, 123
 - de Sobolev, 57
 - duale, 14
 - hyperbolique, 81
 - höldérienne, 43
 - riemannienne, 132
 - uniforme, 13
 - équivalentes, 115, 117
- opérateur de transfert, 24, 62
- orbite, 7
 - complète, 7
 - positive, 6, 7
 - périodique, 7
- ordre de Sharkovsky, 10
- paramétrage local, 5
- partie
 - ϵ -dense, 167
 - fractionnaire, 189
 - hyperbolique, 118
 - localement maximale, 149
 - minimale, 22
 - ϵ -séparée, 167
- partition
 - de Markov, 198
 - image réciproque, 172
 - jointe, 171
 - mesurable, 171
 - génératrice, 174
 - moins fine, 172
- période, 7
- périodique, 7
 - attractif, 7
 - répulsif, 8
- phénomène de Mautner, 75
- pistage, 135, 138
 - linéaire, 152
- pièce basique, 151
- plan hyperbolique réel, 81
- point
 - à l'infini, 86
 - errant, 19
 - extremal, 13
 - fixe hyperbolique, 108
 - homocline, 154
 - récurrent
 - négativement, 18
 - positivement, 18
- positivement minimal, 21
- principe variationnel, 174
- produit local, 149, 196
- projection
 - canonique, 50
 - stéréographique, 10
- propre, 14
- préserver la mesure, 5
- pseudo-orbite, 135, 138, 152
- rayon géodésique, 85
- rayon spectral, 154, 181
- recouvrement
 - générateur, 170
 - joint, 166
 - moins fin, 167
- rectangle, 197
 - propre, 198
- récurrent
 - négativement, 18
 - positivement, 18
- relèvement, 111
- relevé, 34
- représentation
 - quasi-régulière, 73
 - unitaire, 73
- répulsif, 8
- réseau, 72, 80

- cocompact, 72, 80
- uniforme, 72, 80
- rotation, 8
- section, 116
 - ensembliste, 116
- semi-conjugaison, 6
 - ensembliste, 6
- semi-conjugués
 - ensemblissement, 6
- séparable, 17
- séparante, 51
- ϵ -séparé, 167
- sous-décalage de type fini
 - bilatère, 180
 - unilatère, 180
- sous-espace
 - instable, 101
 - stable, 101
- sous-fibré, 115
 - instable, 118, 123
 - neutre, 123
 - stable, 118, 123
- sous-rectangle
 - instable, 198
 - stable, 198
- spectre, 103, 154, 156
- stable, 101, 118, 123
- structurellement stable, 152
 - linéairement, 151
- suite admissible, 205
- suite de Farey, 189
- support, 17
- système de Bernoulli
 - bilatère, 8
 - unilatère, 8
- système de Markov, 180
- système dynamique, 5
 - à temps continu, 5
 - à temps discret, 5
 - conjugué, 6
 - différentiable, 5
 - de classe C^k , 5
 - invertible, 5
 - mesurable, 5
 - mesuré, 5
 - Nord-Sud, 10
 - structurellement stable, 152
 - topologique, 5
 - expansif, 134
 - non errant, 19
- théorème
 - d'Anosov, 113
 - d'Arzela-Ascoli, 92
 - d'inversion locale, 163
 - de Baire, 20
 - de Banach, 154
 - de Banach-Alaoglu, 15
 - de Birkhoff, 16
 - de Bowen, 202
 - de Carathéodory, 184
 - de Dani-Furstenberg, 91
 - de Denjoy, 38
 - de dualité de Riesz, 78
 - de Furstenberg-Weiss, 25
 - de Grobman-Hartman, 108
 - de Howe-Moore, 77
 - de Kolmogorov-Sinaï, 174
 - de Krein-Milman, 15
 - de Krylov-Bogolyubov, 13
 - de Perron-Frobenius, 181
 - de récurrence
 - de Birkhoff, 23
 - de Poincaré, 23
 - de représentation de Riesz, 14
 - de Sard, 153
 - de semi-conjugaison de Poincaré, 39
 - de Sharkovsky, 10
 - de Stone-Weierstrass, 51
 - de van der Waerden, 28
 - du point fixe de Banach, 49
 - du principe variationnel, 174
 - ergodique de Birkhoff, 16
- topologie
 - C^1 , 113, 131
 - C^k , 131
 - faible-étoile, 14
 - vague, 14
- tore, 50
- totalemt discontinu, 9
- transformation, 5
- transitif, 20
 - positivement, 20
 - topologiquement, 20
- translation, 50
 - à droite, 50
 - à gauche, 50
- trivialisation locale, 114
- uniformément continue, 105
- unimodulaire, 70
- valeur
 - critique, 153
 - régulière, 153

variation, 13
 totale, 14
variation bornée, 38
variété
 instable, 147, 196
 locale, 143, 195
 stable, 147, 196
 locale, 143, 195
volume hyperbolique, 81

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.
- [AP] D. K. Arrowsmith and C. M. Place. *An introduction to Dynamical systems*. Cambridge Univ. Press, 1990.
- [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.
- [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.
- [BuzS] J. Buzzi and O. Sarig. *Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps*. Erg. Theo. Dyn. Sys. **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.
- [DS] T.-C. Dinh and N. Sibony. *Equidistribution problems in complex dynamics of higher dimension*. Internat. J. Math. **28** (2017) 1750057.
- [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.
- [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.
- [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.
- [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.
- [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.
- [Paul1] F. Paulin. *De la géométrie et de la dynamique de $SL_n(\mathbb{R})$ et $SL_n(\mathbb{R})$* . 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.

- [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.
- [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.
- [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.
- [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.
- [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.
- [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.
- [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.
- [Rud] W. Rudin. *Functional Analysis*. 2nd ed., McGraw-Hill, 1991.
- [Sar1] O. Sarig. *Existence of Gibbs measures for countable Markov shifts*. Proc. Amer. Math. Soc. **131** (2003) 1751–1758.
- [Sar2] O. Sarig. *Subexponential decay of correlations*. Invent. Math. **150** (2002) 629–653.
- [Sar3] O. Sarig. *Thermodynamic formalism for countable Markov shifts*. Erg. Theo. Dyn. Syst. **19** (1999) 1565–1593.
- [Sar4] 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.
- [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.
- [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