Petite introduction aux systèmes dynamiques

Daniel Massart

Février, 2021

Chapter 1

Introduction

Définition 1.0.1. *Pour nous, un système dynamique sera la donnée*

- *d'un espace métrique* (X, d)
- *(cas du temps discret) d'une transformation continue* $T : X \longrightarrow X$
- *(cas du temps continu) d'un morphisme de* T *de* $(\mathbb{R}, +)$ *dans le groupe des homéomorphismes de* X.

On imagine que X représente "les états du système" (par exemple, les positions de toutes les planètes du système solaire) et T représente la loi d'évolution du système (ce qui nous dit la position à l'instant $t+1$, connaissant la position à l'instant t). Par exemple, dans le cas continu, la loi d'évolution peut être une équation différentielle. Par exemple de l'exemple, l'équation de Newton pour les planètes.

A parte : j'aime bien parler de mécanique céleste pour illustrer ce cours, et en fait la mécanique céleste est un peu à l'origine de toute la théorie des systèmes dynamiques. Mais pour des raisons techniques, je ne démontrerai aucun résultat de mécanique céleste, parce que de toutes les petites hypothèses qu'on aime bien faire pour se simplifier la vie (X est compact, T est continue...), aucune n'est vérifiée dans un vrai système solaire: la mécanique céleste est une maîtresse exigeante.

Dans ce cours je considèrerai souvent un modèle jouet de mécanique céleste, le pendule : une masse ponctuelle est suspendue au bout d'une tige rigide, l'autre extrémité fixée au mur.

Exercice 1. *Quel est l'espace* X *dans ce cas ? et quelle est la loi d'évolution?*

On a bien sûr envie de répondre que la position du pendule est entièrement décrite par l'angle de la tige avec la verticale, et que X est donc le cercle $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Mais cette réponse n'est pas satisfaisante parce que connaître la position du pendule ne suffit pas à déterminer sa trajectoire, il faut aussi connaître sa vitesse. L'espace X est donc l'ensemble des couples (x, v) (position, vitesse), où la position est dans le cercle, et la vitesse est dans $\mathbb{R}$: c'est un cylindre (infini). Si vous avez fait géométrie différentielle, c'est le fibré tangent au cercle. Quant à la loi d'évolution, elle est donnée par l'équation de Newton : force = masse fois accélération (on va supposer que la masse vaut 1, et négliger les frottements)

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= \sin x \end{aligned}$$

Questions : et si on incline le mur ? de vertical, jusqu'à horizontal ? et si on rajoute les frottements ?

Vous avez déjà fait des systèmes dynamiques sans le savoir : dans les exercices bébêtes de L1 du type "$u_{n+1} = f(u_n)$, calculer $\lim u_n$". En fait ces exercices ne sont pas si bêtes que ça et on y trouve plusieurs phénomènes importants (cf. chapitre 2).

Dans les exemples, X sera souvent compact : l'intervalle $[0, 1]$, le cercle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, ou l'ensemble de suites $\{0, 1\}^{\mathbb{N}}$, qui est compact par le théorème de Tykhonov, et T sera souvent, sur le cercle,

-une rotation d'angle θ : $\mathbb{T} \longrightarrow \mathbb{T}$, $x \longmapsto x + \theta \mod 1$,

-ou le doublement : $x \longmapsto 2x \mod 1$;

ou, sur les suites, le décalage : $\{0, 1\}^{\mathbb{N}} \longrightarrow \{0, 1\}^{\mathbb{N}}$, $(u_n)_{n\in\mathbb{N}} \longmapsto (u_{n+1})_{n\in\mathbb{N}}$.

Un autre exemple qui nous accompagnera tout au long de ce cours est la *famille quadratique* : les applications de $[0, 1]$ dans lui-même, de la forme $f_\lambda(x) = \lambda x(1 - x)$.

Définition 1.0.2. *L'orbite (ou trajectoire) d'un point $x \in X$ est l'ensemble de ses itérés : $\{T^n(x) : n \in \mathbb{N}\}$ (temps discret) ou $\{T^t(x) : t \in \mathbb{R}\}$ (temps continu), n (ou t) jouant le rôle du temps.*

Exercice 2. *Trouver des orbites pour les exemples considérés ci-dessus.*

La théorie des systèmes dynamiques étudie les propriétés asymptotiques des orbites (les propriétés vérifiées à très long terme, i.e.quand le temps t tend vers l'infini). Ce point de vue a été introduit par Poincaré au début du XXème siècle, quand il s'est aperçu qu'essayer de résoudre (avec des formules) les équations de la mécanique céleste est une entreprise désespérée, dès qu'il y a au moins deux planètes (ce qui n'empêche pas d'avoir d'excellentes approximations numériques : on peut prédire une éclipse dans 50 ans à la seconde près !). Il s'est dit alors qu'on pouvait au moins répondre à des questions qualitatives : est-il possible qu'un jour Saturne s'échappe du sytème solaire ? que Mars soit capturée en orbite de Jupiter ?

Exercice 3. *Revenons au cas du pendule : savez-vous résoudre l'équation différentielle ? Pouvez-vous quand même donner une description qualitative satisfaisante des orbites ?*

Maintenant imaginons un système un peu plus complexe : un pendule double (au lieu d'une masse, vous mettez un autre pendule). Pouvez-vous encore décrire le système ?

Exercice 4. *$X = [0,1]$, $T(x) = x^2$ (c'est le cas discret ou continu ?). Décrire toutes les orbites.*

Dans cet exemple, on considère qu'on a tout compris au système dynamique (X, T) une fois qu'on a dit "il y a un équilibre stable et un équilibre instable, et l'équilibre stable attire tout le monde, sauf l'équilibre instable". L'attraction est un exemple de propriété asymptotique, on ne cherche pas à connaître exactement la trajectoire d'un point, on veut juste connaître sa limite.

On vient de voir un exemple d'un système particulièrement prévisible : à partir de très peu d'information sur la position initiale (juste $x \neq 1$), on obtient une information asymptotique complète (l'orbite de x devient rapidement indistinguable de celle de 0). Même si on n'a pas encore défini ce qu'est un système chaotique, on a clairement envie de dire que ce système n'est pas chaotique ! En fait le but de ce cours sera de donner une définition à peu près satisfaisante du chaos. La notion intuitive est illustrée par l'effet papillon : le battement d'aile d'un papillon au Brésil peut provoquer une tempête au Texas six mois plus tard. En termes mathématiques, cela signifie qu'une information initiale, si précise soit-elle (être capable de mesurer l'état de l'atmosphère avec une telle précision qu'on pourrait détecter quelque chose d'à peine plus gros qu'un battement d'aile de papillon !) ne donne aucune information asymptotique. Notons qu'il y a quand même de l'imprévisibilité dans notre exemple bébête : des points arbitrairement proches de 1 ont des orbites radicalement différentes de celle de 1. Cela signifie que la notion de chaos est plus délicate à définir qu'il n'y paraît.

Exercice 5. *Y a-t-il de l'imprévisibilité dans les rotations ? dans le doublement ?*

Référence à peu près unique, tant il y a tout dans ce bouquin : Introduction to the modern theory of dynamical systems, par Katok et Hasselblatt. Je recommande également la version "pour enfants" (lire : undergraduates) : A first course in dynamics, des mêmes auteurs. Lire absolument l'introduction. Ne pas oublier de regarder sur Youtube la série "Chaos", d'Etienne Ghys et Jos Leys (ça commence très bas, pour rejoindre le niveau de ce cours à l'épisode 4, et finir assez fort).

Remerciements à Zakaria Baammi, Zakaria Brahimi, Maxime Maillard, Pablo Montealegre, et Cassandre Saiz, pour leur relecture attentive des quatre premiers chapitres.

Chapter 2

Attracteurs, sous-systèmes

2.1 introduction

Une stratégie, face à un système dynamique incompréhensible, est de diviser pour mieux régner : essayer de trouver des sous-systèmes compréhensibles. C'est ce que nous allons faire dans ce chapitre. Un sous système de (X, T) est un couple (Y, T_Y), où Y est une partie fermée (donc compacte si X est compact) et T-invariante de X, et T_Y est la restriction de T à Y (dans la pratique on omettra souvent le Y en indice). Comme une partie ne peut pas être plus compliquée que le tout, il est intéressant, pour comprendre un système dynamique, de chercher d'abord des sous-systèmes simples. Il arrive qu'il n'en existe pas, c'est le cas, par exemple, lorsque toute orbite de T est dense dans X. Cette démarche se retrouve en arithmétique, où on cherche à décomposer un entier en facteurs premiers, et en algèbre, où pour comprendre un groupe, on cherche d'abord ses sous-groupes (distingués). Un système où toute orbite est dense sera pour nous l'équivalent d'un nombre premier, ou d'un groupe simple.

On a bien envie de dire qu'un système qui se décompose en sous-systèmes n'est pas trop chaotique, ou du moins, que le chaos est contenu à l'intérieur d'un sous-système. Une de nos premières tâches sera d'arriver à bien définir la notion d'indécomposabilité (analogue à la primalité pour les entiers, ou à la simplicité pour les groupes). On verra qu'en fait il y a plusieurs notions, suivant les outils utilisés : topologie ("toute orbite est dense", cf. chapitre 3), ou théorie de la mesure (chapitre 4) ; c'est une des difficultés des systèmes dynamiques.

Le plus simple sous-système possible est celui qui ne contient qu'un seul point, nécessairement fixe (un point $x \in X$ est fixe par T si $T(x) = x$). Un point fixe est parfois appelé équilibre dans la littérature d'inspiration physique. Le théorème du point fixe pour les applications contractantes des complets

Exercice 6. *Enoncer et démontrer le théorème* est le deuxième théorème de systèmes dynamiques ! Le premier, c'est quand vous avez trouvé la limite de u_n, avec $u_{n+1} = f(u_n)$, au lycée.

Sauriez-vous démontrer la version où X est compact, mais T est seulement contractante au sens où $d(T(x), T(y)) < d(x, y)$ pour tous $x, y \in X$? et (beaucoup plus difficile) celle où X est un compact convexe de $\mathbb{R}^n$, mais où on suppose seulement que T est continue ?

Le deuxième cas le plus simple est celui d'un ensemble invariant fini. Un tel ensemble est formé de points périodiques (ou ultimement périodiques, si T n'est pas inversible), c'est à dire des points $x \in X$, tels qu'il existe $n \in \mathbb{N}$, tel que $T^n(x) = x$. Trouver une orbite périodique d'un système donné est souvent un exploit, salué comme tel par la communauté, en particulier en mécanique céleste. Comprendre les orbites périodiques permet souvent de mieux comprendre le système global (intuition originellement due à Poincaré), comme par exemple dans le théorème suivant.

Pour énoncer le théorème il faut d'abord définir une relation d'ordre bizarre sur $\mathbb{N}$, dite ordre de Sharkovsky et notée $\prec$: on dit que $2^k(2p+1) \prec 2^{k'}(2p'+1)$ si

- $p, p' > 0$ et $k < k'$, ou
- ou $p, p' > 0$, $k = k'$, et $p < p'$ (ordre lexicographique sur la paire (k, p), quand $p > 0$)
- $p > 0$ et $p' = 0$ (les puissances de 2 sont tout en haut de la pile)
- $p = p' = 0$ et $k > k'$ (attention à l'inversion de l'ordre sur les puissances de 2).

Exercice 7. *Montrer que l'ordre de Sharkovsky possède un maximum et un minimum (contrairement à l'ordre naturel). Trouver la relation d'ordre entre 1001 et 1111. Montrer que tout élément sauf* 1 *admet un successeur, et que tout élément admet un prédécesseur, sauf les multiples de* 3.

Théorème 2.1.1 (Sharkovsky, 1964)**.** *Si une application continue* $T : [0, 1] \longrightarrow [0, 1]$ *admet une orbite de période* n*, alors elle admet des orbites périodiques de toutes périodes* k*, avec* $n \prec k$*.*

Ce théorème est souvent énoncé sous la forme "période trois implique chaos", même si le mot chaos est ici employé dans un sens faible (il veut juste dire que la dynamique est compliquée). Il peut faire l'objet d'un exposé.

2.2 Attracteurs

La connaissance d'un sous-système (Y, T_Y) permet parfois de comprendre la dynamique, pas seulement sur le support Y du sous-système, mais au voisinage de Y.

Définition 2.2.1. *On dit qu'un point fixe x de T est attractif (on parle aussi d'équilibre stable) s'il existe un voisinage U de x dans X, tel que pour tout $y \in U$, $T^n(y)$ tend vers x quand n tend vers $+\infty$. Le bassin d'attraction de x est alors l'ensemble des $y \in X$ tels que $T^n(y)$ tend vers x quand n tend vers $+\infty$.*

De même on dit qu'un point fixe est répulsif (équilibre instable) s'il existe un voisinage U de x dans X, tel que pour tout $y \in U$, il existe $N \in \mathbb{N}$, tel que $\forall n \geq N$, $T^n(y) \notin U$.

Exercice 8. *Les points fixes de l'exercice 4 sont-ils attractifs ? répulsifs ? Et les équilibres du pendule ?*

Un point périodique est juste un point fixe de T^n pour un certain n, donc on peut définir de la même façon une orbite périodique attractive ou répulsive.

Exercice 9. *Trouver des orbites pérodiques pour les rotations, et le doublement. Sont-elles attractives ? répulsives ?*

En fait on peut même définir un ensemble attractif quelconque :

Définition 2.2.2. *On appelle attracteur une partie fermée F de X, telle qu'il existe un voisinage U de F dans X, tel que $T(U) \subset U$, et*

$$F = \bigcap_{n \in \mathbb{N}} T^n(U).$$

On verra au chapitre comment exprimer cette définition en termes d'ensemble ω-limite.

On verra aussi, au paragraphe suivant, qu'en dimension (de X) 2 les attracteurs ne sont pas très variés. En revanche dès la dimension 3 il peut arriver des choses incroyables, par exemple (exposé possible) l'attracteur de Lorentz (météorologue américain, inventeur de l'effet papillon).

Quand on a identifié un attracteur, on a ramené l'étude de la dynamique au voisinage (plus exactement, dans le bassin d'attraction) de l'attracteur à l'étude de la dynamique dans l'attracteur. Celle-ci est bien sûr triviale si l'attracteur est un point fixe ou une orbite périodique, mais elle peut être très compliquée en général (cf. chapitre suivant).

Exercice 10. *Montrer qu'un attracteur (pour T) est toujours stable par T.*

Vocabulaire Dans le cas où T est inversible, on dit qu'un point $x \in X$ est homocline à un autre point $y \in X$ si $d(T^n(x), T^n(y)) \longrightarrow 0$ quand $n \to \pm\infty$ (n ne peut tendre vers $-\infty$ que dans le cas où T est inversible). On dit que x est hétérocline à deux points y, z si $d(T^n(x), T^n(y)) \longrightarrow 0$ quand $n \to +\infty$ et $d(T^n(x), T^n(z)) \longrightarrow 0$ quand $n \to -\infty$.

2.3 La famille quadratique

Il s'agit de la famille d'application $f_\lambda : [0,1] \to [0,1]$, $x \mapsto \lambda x(1-x)$, le paramètre λ variant entre 0 et 4 (pour garantir que l'application est bien à valeurs dans $[0,1]$). Que peut-il y avoir d'intéressant dans un truc aussi simple ? Apparemment pas grand chose :

Exercice 11. *Décrire complètement la dynamique pour $\lambda \in [0,1]$. Que se passe-t-il pour $\lambda = 1$?*

La réponse attendue est que le point fixe est moins attractif pour $\lambda = 1$: la convergence est moins rapide.

Que se passe-t-il ensuite ? deux choses : $\lambda > 1$ donc le point fixe à l'origine devient répulsif, et d'autre part, un autre point fixe apparaît, en $x = (\lambda - 1)/\lambda$, où la dérivée vaut $(2-\lambda)/\lambda$, qui est entre 0 et 1, strictement, pour λ entre 1 et 2, donc le nouveau point fixe est attractif, et la dynamique n'est toujours pas très mystérieuse : tout le monde est attiré par le nouveau point fixe, sauf 0 qui boude dans son coin (dynamiquement, c'est pareil que l'application de l'exercice 4). Notez que c'est comme si le point fixe attractif 0 s'était dédoublé pour donner naissance à une paire de points fixes, l'un répulsif, et l'autre attractif. Voir l'excellente video de Veritasium, grotesquement intitulée "cette équation va changer la façon dont vous voyez le monde", pour des animations.

Que se passe-t-il ensuite ? Ensuite la dérivée en le nouveau point fixe devient négative (le point fixe a passé la bosse), mais reste < 1 en valeur absolue, donc le point fixe reste attractif, simplement les orbites dessinent des escargots, ou des toiles d'araignées, autour du point fixe.

Les choses deviennent intéressantes à partir de $\lambda > 3$: la dérivée de f_λ passe au dessus de 1 en valeur absolue, le point fixe devient répulsif. Mais dans ce cas, où vont les orbites ? Pour comprendre ce qui se passe, regardons l'application $f_\lambda^2 = f_\lambda \circ f_\lambda$ (le carré est au sens de l'itération, non de la multiplication). On voit que pour $\lambda > 3$, il apparaît un point fixe de f_λ^2 qui n'est pas point fixe de f_λ, c'est donc une orbite périodique de f_λ, de période 2. Tant que λ n'est pas trop loin de 3, la dérivée de f_λ^2 au point fixe est < 1, l'orbite périodique est donc attractive. La dynamique est juste un peu plus compliquée, on a les deux points fixes répulsifs, et tout le reste est attiré par l'orbite périodique.

Puis, quand on continue à augmenter λ, il vient un moment où la dérivée de f_λ^2 au point fixe devient > 1 en valeur absolue, l'orbite périodique devient répulsive. Il convient alors de regarder f_λ^4 : on constate qu'au moment où l'orbite 2-périodique devient répulsive, il apparaît, au voisinage de chaque point fixe de f_λ^2, une paire de points fixes de f_λ^4. C'est donc une orbite de période 4 de f_λ qui apparaît, attractive dans un premier temps, puis répulsive, etc... à chaque fois qu'une orbite de période 2^n devient répulsive, elle se dédouble en quelque sorte, et il apparaît une orbite de période 2^{n+1}.

Vous voyez venir l'ordre de Sharkovsky ? Et du coup, on se pose la question : est-ce qu'on arrive au bout (puisque l'ordre de Sharkovsky, contrairement à l'ordre naturel, possède un minimum) ? Pour cela on regarde f_λ^3, et on constate que vers $\lambda = 3.82$, il apparaît une orbite de période 3.

La suite sort du cadre de ce chapitre, mais peut faire l'objet d'un exposé (pour lequel on consultera avec profit l'excellent texte de vulgarisation de Daniel Perrin).

2.4 Champs de vecteurs dans le plan

Un champ de vecteurs sur le plan est une application continue $X : \mathbb{R}^2 \to \mathbb{R}^2$. Il faut imaginer qu'en chaque $x \in \mathbb{R}^2$, vu comme un point, on attache $X(x)$, vu comme un vecteur ; on a donc un champ de vecteurs, comme on a un champ de blé, les points sont les racines, et les vecteurs sont les épis. Un champ de vecteurs définit une équation différentielle : $f'(t) = X(f(t))$, où la fonction inconnue f est définie sur un intervalle de $\mathbb{R}$, à valeurs dans $\mathbb{R}^2$ (il faut la voir comme une courbe paramétrée dans le plan, dont le vecteur-vitesse en chaque point est $X(f(t))$). Sous des hypothèses bénignes sur le champ de vecteur X (nous le supposerons C^1), l'existence de solutions maximales définies sur $\mathbb{R}$ est garantie par le théorème de Cauchy-Lipschitz. Pour tout $x \in \mathbb{R}^2$, il existe donc $f_x : \mathbb{R} \to \mathbb{R}^2$ qui est solution de l'équation différentielle, et telle que $f_x(0) = x$. On appelle flot de X l'application $\phi_t^X : \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}^2, (x,t) \mapsto f_x(t)$. Si on pense aux trajectoires comme aux lignes de courant d'un écoulement de fluide dans le plan, cette application consiste à suivre le courant (la trajectoire du point x) pendant un temps t, d'où son nom de flot. Il est facile de vérifier que l'application $t \mapsto \phi_t^X$ est un morphisme de $(\mathbb{R}, +)$ dans le groupe des homéomorphismes de $\mathbb{R}^2$, c'est donc un système dynamique (cas du temps continu), et nous allons nous intéresser à ses attracteurs.

Théorème 2.4.1 (Poincaré-Bendixson, ca. 1880)**.** *Les seuls attracteurs possibles pour un champ de vecteurs dans le plan sont un point fixe, une orbite périodique, ou un cycle (pas nécessairement fini) de points fixes joints par des orbites homoclines.*

Du point de vue de la dynamique, c'est un résultat triste : il dit que les champs de vecteurs dans le plan, ce n'est pas très intéressant. Plus précisément, il dit que la topologie du plan (ou de la sphère) est trop simple pour porter des dynamiques continues compliquées. La précision "continues" est nécessaire, parce que vous avez tous vu un ensemble de Mandelbrot, aucune personne de bon sens ne prétendrait que ce n'est pas compliqué. La précision "plan ou sphère" est nécessaire aussi, parce que nous verrons des exemples de dynamiques plus compliquées sur des surfaces de genre plus grand.

La preuve est de nature topologique (l'outil essentiel est le thm de Jordan sur les courbes planes). L'idée est très bien expliquée à l'épisode 4 de Chaos.

Exposé : sur le théorème de Jordan ? package avec Poincaré-Bendixson ?

Par contre ce théorème est souvent utile dans les applications. Par exemple c'est la raison pour laquelle la dynamique du pendule est aussi simple (évidemment, ça ne marche plus pour le pendule double).

Pour autant on est loin de tout comprendre à la dynamique plane, cf. le célèbre XVIème problème de Hilbert. Et rassurez-vous, dès qu'on passe à la dimension 3, les choses se compliquent, et il n'existe aucune généralisation, même conjecturale, du théorème de Poincaré-Bendixson aux dimensions > 2. En gros, n'importe quoi peut être un attracteur d'un système dynamique, aucun espoir de classifier.

Esquisse de preuve (la preuve complète peut faire l'objet d'un exposé)

Observons que si toutes les orbites partent à l'infini, il n'y a pas d'attracteur, donc le théorème est démontré. On pourrait dire aussi que dans ce cas, l'infini est point fixe attractif ; nous verrons au chapitre suivant comment formaliser cette observation, à l'aide de la compactification d'Alexandrov. Supposons donc qu'il existe une orbite qui ne part pas à l'infini :

$$\exists x \in \mathbb{R}^2, \exists A > 0, \forall T \in R, \exists t > T, \|f_x(t)\| \leq A,$$

autrement dit, l'orbite de x repasse une infinité de fois dans la boule centrée en l'origine, de rayon A. Mais les boules sont compactes, l'orbite de x admet donc un point d'accumulation : il existe une suite t_n qui tend vers l'infini, et y dans la boule de rayon A, tels que $f_x(t_n) \to y$ quand n tend vers l'infini. Là il y a deux cas à distinguer, suivant que $X(y) = 0$ ou non. Nous allons traiter le cas le plus facile, celui où $X(y) \neq 0$.

Premier cas : $X(y) \neq 0$. Alors par continuité, il existe une boule centrée en y dans laquelle le champs X ne s'annule pas. Il existe donc un (petit) segment T transverse au flot centré en y. On considère la suite des temps de retour s_n de l'orbite f_x sur la transversale T. On fabrique une courbe fermée γ_n dans le plan en recollant le morceau d'orbite $f_x([s_n, s_{n+1},])$ avec une portion du segment T joignant $f_x(s_n)$ à $f_x(s_{n+1})$. Par le théorème de Jordan, la courbe γ_n sépare le plan en deux composantes connexes. L'une des deux est un piège : aucune orbite ne peut s'en échapper.On fabrique ainsi (en faisant tendre n vers l'infini) une suite de pièges emboîtés. On prend l'intersection, qui est non vide par le théorème des compacts emboîtés (exercice pour ceux qui ne le connaissent pas), c'est un compact invariant par le flot, son bord est invariant (chaque temps du flot est un homéomorphisme, donc il préserve la propriété d'être sur le bord), donc c'est nécessairement une orbite périodique, s'il ne contient pas de point fixe, ou une chaîne homocline, s'il contient un ou plusieurs points fixes.

Chapter 3

Conjugaison topologique et stabilité structurelle

3.1 Dynamique linéaire dans $\mathbb{R}$

Commençons par le cas le plus simple : les endomorphismes linéaires de $\mathbb{R}$. Rappelons que tout endomorphisme linéaire de $\mathbb{R}$ est une homothétie, de la forme $T_\lambda : x \mapsto \lambda x$.

Exercice 12 (à faire en classe)**.** *Décrire la dynamique de toutes les applications linéaires $\mathbb{R} \longrightarrow \mathbb{R}$:*

1. *si $|\lambda| < 1$, l'origine est un point fixe attractif, tout point de $\mathbb{R}$ est dans le bassin d'attraction de l'origine*
2. *si $|\lambda| > 1$, l'origine est point fixe répulsif, l'item 1 s'applique à T^{-1}*
3. *si $|\lambda| = 1$, T est périodique : $T = id$ ou $T^2 = id$.*

Si vous avez l'impression que, sauf dans le cas exceptionnel $|\lambda| = 1$, du point de vue dynamique, les endomorphismes linéaires de $\mathbb{R}$ se ressemblent tous un peu (ou du moins, les T_λ, $|\lambda| < 1$ se ressemblent entre eux, et de même les T_λ, $|\lambda| > 1$), vous avez raison. Peut-on tirer une conséquence pertinente de cette observation un peu morose ?

3.1.1 Conjugaison et semi-conjugaison

Comment définir de façon rigoureuse le fait "d'avoir la même dynamique"? la solution est dans une notion que vous avez vu en L1 : le changement de variable. Vous observez un système dynamique, puis vous changez de lunettes et vous voyez autre chose : vous avez envie de dire que les deux sont équivalents modulo changement de lunettes.

Plus précisément :

Définition 3.1.1. *Si $f : X \longrightarrow Y$ est un homéomorphisme, alors $f \circ T \circ f^{-1}$ définit un système dynamique sur Y, qui est dit topologiquement conjugué à T.*

Du point du vue dynamique, T et $f \circ T \circ f^{-1}$ sont essentiellement équivalents : à tout point fixe (resp. périodique, resp. attracteur, etc...) de T correspond, par f, un point fixe (resp. périodique, resp. attracteur, etc...) de f.

Une version plus faible et non-symétrique de la relation de conjugaison nous sera souvent utile :

Définition 3.1.2. *Si (X, T) et (Y, S) sont deux systèmes dynamiques, et $f : X \longrightarrow Y$ est une application continue, telle que $f \circ T = S \circ f$, alors on dit que f est une semi-conjugaison de T sur S, ou que S est un facteur de T.*

Le mot facteur vient du fait que S est en quelque sorte contenu dans T, comme un facteur premier d'un entier est contenu dans cet entier (pas très sûr de moi là-dessus). L'image par f d'un point fixe (resp. périodique, resp. attracteur, etc...) de T est un un point fixe (resp. périodique, resp. attracteur, etc...) de S, mais la correspondance peut n'être pas univoque, puisque f n'est pas supposée injective.

Laissons de côté un moment la dynamique linéaire pour donner un exemple important de semi-conjugaison. C'est aussi le moment d'introduire une idée importante en systèmes dynamiques : le codage des orbites. Prenons $X = \mathbb{R}/\mathbb{Z}$, T est le doublement, et notons E_0 le demi-cercle ouvert "supérieur" (image dans X de l'intervalle $]0, 1/2[$), et E_1 le demi-cercle "inférieur" (image dans X de l'intervalle $]1/2, 1[$). Notons $\mathcal{O}$ l'orbite négative de 0 (la réunion, sur $n \in \mathbb{N}$, des $T^{-n}(\{0\})$). Ce n'est rien d'autre que l'ensemble des rationnels dyadiques de $[0, 1[$. Pour $x \in X$, on définit une suite $(s(x)_n)_{n\in\mathbb{N}}$ par $s(x)_i = 0$ si $T^i(x) \in E_0$, $s(x)_i = 1$ si $T^i(x) \in E_1$. On appelle itinéraire de x la suite $(s(x)_n)_{n\in\mathbb{N}}$.

Exercice 13.

1. *Montrer que l'itinéraire de x est égal au développement binaire de x, si x n'est pas un rationnel dyadique, et est égal au développement binaire de x qui finit par des 0, si x est un rationnel dyadique.*

2. *Montrer que $s(Tx)$ est l'image de la suite $s(x)$ par le décalage.*

3. *Montrer que l'application "itinéraire" $s : x \mapsto s(x)$ est un homéomorphisme de $X \setminus \mathcal{O}$ sur $\{0, 1\}^{\mathbb{N}}$, privé des suites constantes à partir d'un certain rang, muni de la topologie produit.*

4. *Montrer que s^{-1} est une semi-conjugaison du décalage sur le doublement. Montrer que s^{-1} est injective excepté sur l'ensemble des suites constantes à partir d'un certain rang.*

Le fait que s soit injective en dehors d'un ensemble négligeable nous dit que du point de vue de la mesure, s est en fait une conjugaison, donc toutes les propriétés de théorie de la mesure (nous en verrons aux chapitres suivants) valables pour le doublement, le sont aussi pour le décalage (et réciproquement).

A l'époque du boom des systèmes dynamiques (les années 60), les dynamiciens avaient l'ambition de classifier les systèmes dynamiques à conjugaison topologique près, comme les topologues espéraient classifier les variétés à homéomorphisme près, les théoriciens des groupes, les groupes à isomorphisme près, etc... Depuis, on a compris que ça risque d'être difficile ; dans ce cours nous verrons quelques classes particulières de systèmes qu'on peut classifier : outre ce chapitre, nous traiterons les applications expansives du cercle, et les homéomorphismes du cercle. Et c'est déjà pas mal de boulot.

Le choix de la relation d'équivalence n'est pas anodin :

Exercice 14. *Montrer que deux endomorphisme linéaires de $\mathbb{R}$ distincts ne sont jamais conjugués par une application linéaire.*

La solution de l'exercice suivant va nous montrer qu'on ne peut pas tellement faire mieux que la conjugaison topologique : par exemple, ça ne marcherait pas en remplaçant topologique par différentiable.

Exercice 15. *Montrer que deux endomorphisme linéaires de $\mathbb{R}$ T_λ et T_μ avec $\lambda, \mu > 1$ (resp. $1 > \lambda, \mu > 0$) sont topologiquement conjugués.*

Solution On vient de voir que ça ne peut pas marcher avec une conjugaison linéaire ; quelle est la conjugaison la plus simple après linéaire ? un polynôme ? et quels sont les polynômes les plus simples : mais oui, les monômes. Cherchez une condition sur $a \in \mathbb{R}$ pour que T_λ et T_μ soient conjugués par x^a. Oui je sais, vous allez me dire "ça veut dire quoi x^a si a n'est pas un entier ?". Ici on veut un homéomorphisme de $\mathbb{R}$, donc pour nous, $x \mapsto x^a$, ce sera la fonction impaire qui vaut x^a si $x > 0$. Vous devez trouver $a = \log \lambda / \log \mu$ (ou $\log \mu / \log \lambda$).

Et là, vous vous dites que ça marcherait même si λ et μ n'étaient pas tous les deux > 1, à ceci près que si $\lambda > 1 > \mu > 0$, a est négatif, et donc la conjugaison n'est pas définie en zéro : elle envoie zéro à l'infini. Qu'à cela ne tienne, on n'a qu'à rajouter l'infini au domaine de définition ? Cette idée de rajouter un point à l'infini n'est pas nouvelle et elle porte un nom :

3.1.2 Compactification d'Alexandrov

Définition 3.1.3. *Soit (X, d) un espace métrique. On appelle compactifié d'Alexandrov de X, et on note $\hat{X}$, la réunion disjointe de X et d'un élément noté ∞, muni de la topologie dont les ouverts sont les ouverts de X, et les parties de la forme $(X \setminus K) \cup \{\infty\}$, où K est un compact de X.*

Exercice 16. *Montrer que le compactifié d'Alexandrov d'un espace métrique X, tel qu'il est décrit ci-dessus, est bien un compact, dans lequel X est dense.*

Exercice 17. *Montrer que le compactifié d'Alexandrov de $\mathbb{R}^n$ est homéomorphe à la sphère $\mathbb{S}^n$. Montrer que tout endomorphisme linéaire de $\mathbb{R}^n$ se prolonge par continuité à la sphère, en décidant que l'infini est fixé.*

Appelons $\hat{T}_\lambda$ le prolongement de T_λ à $\hat{\mathbb{R}}$:

$$\begin{array}{rcl} \bar{T}_\lambda : \hat{\mathbb{R}} & \longrightarrow & \hat{\mathbb{R}} \\ x & \longmapsto & \lambda x \text{ si } x \in \mathbb{R} \\ & & \infty \text{ si } x = \infty. \end{array}$$

Exercice 18. *Montrer que $\hat{T}_\lambda$ est un homéomorphisme de $\hat{\mathbb{R}}$.*

indication : une bijection continue d'un compact est un homéomorphisme. Il suffit de vérifier que $\hat{T}_\lambda$ est continue en l'infini.

On continue, par abus de language, à appeler $\hat{T}_\lambda$ endomorphisme linéaire.

De même on étend $x \mapsto x^a$ à $\hat{\mathbb{R}}$, de la façon suivante :

$$\begin{array}{rcl} \bar{\mathbb{R}} & \longrightarrow & \bar{\mathbb{R}} \\ x & \longmapsto & x^a \text{ si } x \in \mathbb{R} \\ & & \infty \text{ si } a > 0 \text{ et } x = \infty \\ & & \infty \text{ si } a < 0 \text{ et } x = 0 \\ & & 0 \text{ si } a < 0 \text{ et } x = \infty \end{array}$$

Exercice 19. *Montrer que deux endomorphisme linéaires de $\hat{\mathbb{R}}$ $\hat{T}_\lambda$ et $\hat{T}_\mu$ avec $\lambda, \mu > 0$, $\lambda, \mu \neq \pm 1$, sont topologiquement conjugués. Que peut-on dire quand $\lambda, \mu = \pm 1$?*

3.1.3 Stabilité structurelle

On a vu que deux endomorphismes linéaires quelconque (avec $\lambda, \mu > 0$ quand même) ont des dynamiques identiques. C'est très bien mais supposons (c'est toujours le cas en pratique) que nous ne disposons que d'instruments de mesure imparfaits pour observer les systèmes. Comment savoir si T est linéaire ? Pourrait-on, malgré l'imperfection de nos observations, dire quelque chose sur la dynamique de T ? Ceci nous amène à la notion de stabilité structurelle

Informellement, on dit qu'un système (X, T) est structurellement stable si pour tout T' proche de T, T' est topologiquement conjugué à T. Bien sûr il faut définir "proche" : puisque nous travaillons avec T continue, la première idée qui nous vient à l'esprit est de dire que deux systèmes T, T' sont ϵ-proches, pour $\epsilon > 0$, s'ils sont ϵ-proches dans la topologie uniforme, c'est à dire que $\forall x \in X, d(T(x), T'(x)) < \epsilon$. Mais c'est une mauvaise idée, à cause de l'exemple suivant.

Exercice 20. *Montrer que pour tout $\epsilon > 0$, il existe T ϵ-proche de T_2 $(x \mapsto 2x)$ dans la topologie uniforme, qui admet une infinité de points fixes.*

Evidemment un tel T ne peut pas être topologiquement conjugué à T_2, qui n'a que deux points fixes (en comptant l'infini). Cela nous oblige à définir plus finement la notion de proximité. Ce qui suit n'a de sens que si X est une variété différentielle (si vous ne savez pas ce que c'est, ce n'est ni plus ni moins qu'un espace où on peut faire du calcul différentiel, c'est à dire définir une notion de fonction dérivable. Si vous n'avez pas envie de savoir ce que c'est, pensez que X est le cercle). Nous entrons donc dans le sous-domaine des systèmes dynamiques différentiables.

Définition 3.1.4. *Supposons que X est une variété différentielle. Soient $\epsilon > 0$, k un entier naturel, et T_1, T_2 (ici rien à voir avec les T_i de la section précédente) deux applications k fois différentiables de X dans X. On dit que T_1 et T_2 sont ϵ-proches dans la topologie C^k si pour tout entier naturel $i \leq k$, les dérivées i-ièmes de T_1 et T_2 sont ϵ-proches dans la topologie uniforme.*

Remarque 3.1.5. *Dire "T_1 et T_2 sont ϵ-proches dans la topologie uniforme" revient donc à dire "T_1 et T_2 sont ϵ-proches dans la topologie C^0".*

Bien sûr on pourrait travailler aussi avec des systèmes très réguliers, par exemple C^∞ ou analytiques, ou au contraire juste mesurables, et on aurait des notions de proximité différentes. Formellement, on a la

Définition 3.1.6. *Un système dynamique (X, T), X étant une variété différentielle, et T de classe C^1, est structurellement stable s'il existe $\epsilon > 0$ tel que pour tout $T' : X \to X$ avec T et T' ϵ-proches dans la topologie C^1, T' est topologiquement conjugué à T.*

Observez que ce qu'on donne d'un côté, on ne le regagne pas complètement de l'autre : on paye le prix en monnaie C^1, mais on achète un produit C^0. C'est la (dure) vie du dynamicien.

Théorème 3.1.7. *L'endomorphisme linéaire T_λ, avec $\lambda \neq 0, \pm 1$, est structurellement stable.*

Proof. Commençons par le cas où $\lambda > 1$. Soit $\epsilon > 0$ tel que $\lambda > 1 + 8\epsilon$, et soit $T : \mathbb{R} \to \mathbb{R}$, ϵ-proche de T_λ dans la topologie C^1. Il s'agit de résoudre l'équation fonctionnelle

$$T_\lambda \circ h = h \circ T \tag{3.1}$$

avec $h : \mathbb{R} \to \mathbb{R}$ continue et strictement croissante.

Posons $\delta = \epsilon/(\lambda - 1)$, et supposons quitte à rétrécir ϵ que $\delta < 1/8$. Soit $r = (1 - \sqrt{1 - 8\delta})/2$, de sorte que $x^2 \geq x - 2\delta$ pour $0 \leq x \leq r$. Définissons une fonction $\tau : \mathbb{R}_+ \longrightarrow \mathbb{R}$ par

$$\begin{aligned} \tau(x) &= x^2 \text{ si } x \leq r \\ &= x - 2\delta \text{ si } x \geq r \end{aligned}$$

Soit $\mathcal{C}$ l'ensemble des applications continues h de $\mathbb{R}$ dans lui-même telles que $\forall x \in \mathbb{R}, |h(x) - x| \leq \delta$, et $\forall x > y \in \mathbb{R}$, $h(x) - h(y) \geq \tau(x - y)$. Munissons $\mathcal{C}$ de la distance uniforme:

$$d(h_1, h_2) = \max_{x \in \mathbb{R}} |h_1(x) - h_2(x)|.$$

Lemme 3.1.8. *L'ensemble $\mathcal{C}$, muni de la distance uniforme, est complet.*

Proof. Prenons une suite de Cauchy dans $\mathcal{C}$, on commence par montrer qu'elle converge simplement, en utilisant la complétude de $\mathbb{R}$, puis on montre qu'en fait la convergence est uniforme. □

La première condition est là pour garantir que la distance est bien définie, malgré la non-compacité de $\mathbb{R}$, et la seconde condition est là pour garantir que le h que nous allons trouver un peu plus loin (et qui *a priori* est juste une application continue de $\mathbb{R}$ dans $\mathbb{R}$) est bien un homéomorphisme.

L'équation (3.1) revient alors à

$$\forall x \in \mathbb{R}, \lambda h(x) = h(T(x)), \text{ avec } h \in \mathcal{C}.$$

Considérons l'application

$$\begin{aligned} \mathcal{F} : \mathcal{C} &\longrightarrow C^0(\mathbb{R}) \\ h &\longmapsto \frac{1}{\lambda} h \circ T. \end{aligned}$$

Lemme 3.1.9. *L'ensemble $\mathcal{C}$ est invariant par $\mathcal{F}$.*

Proof. : pour montrer que la condition $\forall x \in \mathbb{R}, |h(x) - x| \leq \delta$ est invariante :

$$\left| \frac{1}{\lambda} h \circ T(x) - x \right| = \left| \frac{1}{\lambda} \left(h(T(x)) - T(x) + T(x) - \lambda x \right) \right| \leq \frac{\delta + \epsilon}{\lambda} \leq \delta.$$

Pour montrer que la condition

$$\forall x > y \in \mathbb{R}, h(x) - h(y) \geq \tau(x - y)$$

est invariante, il s'agit de montrer

$$\forall x > y \in \mathbb{R}, h(T(x)) - h(T(y)) \geq \lambda \tau(x - y),$$

et il suffit pour cela de montrer que

$$\forall x > y \in \mathbb{R}, \tau\left(T(x) - T(y)\right) \geq \lambda\tau(x-y).$$

Notons qu'on a, puisque T est ϵ-proche de T_λ dans la topologie C^1, $\forall x > y \in \mathbb{R}$,

$$\begin{aligned} T(x) - T(y) &\geq (\lambda - \epsilon)(x-y) \text{ si } x - y \leq 2 \\ &\geq \lambda(x-y) - 2\epsilon \text{ si } x - y \geq 2. \end{aligned}$$

Premier cas : $x - y \geq 2$. Alors

$$\begin{aligned} \tau\left(T(x) - T(y)\right) &= T(x) - T(y) - 2\delta \geq \lambda(x-y) - 2\epsilon - 2\delta \\ &= \lambda(x - y - 2\delta) = \lambda\tau(x-y). \end{aligned}$$

Deuxième cas : $r \leq x - y \leq 2$. On a

$$\begin{aligned} \tau\left(T(x) - T(y)\right) &= T(x) - T(y) - 2\delta \geq (\lambda - \epsilon)(x-y) - 2\delta \\ &\geq \lambda(x - y - 2\delta) = \lambda\tau(x-y) \end{aligned}$$

puisque $2\delta(\lambda - 1) \geq 2\epsilon \geq \epsilon(x-y)$ par définition de δ, et parce que $x - y \leq 2$.

Troisième cas : $x - y \leq r$, mais $T(x) - T(y) \geq r$. On a

$$\tau\left(T(x) - T(y)\right) = T(x) - T(y) - 2\delta \geq (\lambda - \epsilon)(x-y) - 2\delta$$

et on veut donc montrer que $(\lambda - \epsilon)(x-y) - 2\delta \geq \lambda(x-y)^2$. Puisque $x - y \leq r$, on a $(x-y)^2 - (x-y) + 2\delta \leq 0$, il suffit donc de montrer que

$$\begin{aligned} (\lambda - 1)(x-y)^2 - (\lambda - 1 - \epsilon)(x-y) \leq 0 &\iff \\ (\lambda - 1)(x-y) - (\lambda - 1 - \epsilon) \leq 0 &\iff \\ x - y \leq 1 - \delta & \end{aligned}$$

ce qui est vrai, puisque $x - y \leq r \leq 1/2 \leq 1 - \delta$, sachant que $\delta < 1/2$.

Quatrième cas : $T(x) - T(y) \leq r$.

$$\tau\left(T(x) - T(y)\right) = \left(T(x) - T(y)\right)^2 \geq (\lambda - \epsilon)^2(x-y)^2 \geq \lambda(x-y)^2$$

où la dernière inégalité est due au fait que $\lambda > 1 + 8\epsilon$. $\square$

Ce que nous cherchons, c'est donc un point fixe de $\mathcal{F}$. Maintenant, observons que $\mathcal{F}$ est contractante, d'un facteur $1/\lambda$, pour la distance uniforme :

$$(\forall x \in \mathbb{R}, |h_1(x) - h_2(x)| \leq \epsilon) \Longrightarrow \left(\forall x \in \mathbb{R}, \frac{1}{\lambda}|h_1(T(x)) - h_2(T(x))| \leq \frac{\epsilon}{\lambda}\right).$$

On conclut alors avec l'exercice 6. Ceci achève la preuve du cas où $\lambda > 1$. On traite les autres cas au moyen de l'exercice 19. $\square$

Remarque 3.1.10. *Notons qu'en plus de l'existence, on a obtenu la continuité de la conjugaison (l'application h) vue comme fonction de T, dans la topologie uniforme, puisque δ tend vers zéro quand ϵ tend vers zéro.*

3.2 Dynamique linéaire dans $\mathbb{R}^2$

Dans ce paragraphe, X est $\mathbb{R}^2$, ou $\mathbb{R}^n$, et T est une application linéaire. Classification à conjugaison topologique (pas linéaire !) près :

1. si les deux valeurs propres sont de module < 1 : l'application est contractante (vue comme application de $\mathbb{R}^2$), l'origine est attractive, son bassin est $\mathbb{R}^2$ tout entier, l'infini est répulsif (dynamique nord-sud)

2. si les deux valeurs propres sont de module > 1 : l'infini est attractif, son bassin est $\mathbb{R}^2 \setminus \{0\}$, l'origine est répulsive (dynamique sud-nord)

3. si une valeur propre est de module > 1 et l'autre est de module < 1 : l'origine et l'infini sont ce qu'on appelle des points-selle hyperboliques. L'infini attire tous les points, sauf ceux situés sur le sous-espace propre associé à la valeur propre < 1. Les orbites sont situées sur des hyperboles, d'où le nom.

4. deux valeurs propres complexes conjuguées de module 1 : l'origine et l'infini n'attirent ni ne repoussent personne. Les orbites sont situées sur des ellipses (d'où le nom de point fixe elliptique). Sur chaque ellipse, la dynamique est conjuguée à celle d'une rotation d'angle l'argument de la valeur propre.

5. une valeur propre $= \pm 1$, l'autre > 1. (Quitte à considérer T^2 au lieu de T, on peut supposer que c'est 1). L'infini attire tout le monde, sauf le sous-espace propre associé à ± 1, où tous les points sont fixes ou de période deux

6. 1 est valeur propre double, mais T n'est pas l'identité. L'infini attire tout le monde sauf le sous-espace propre (mais moins fort, en un certain sens à préciser, que dans le cas hyperbolique), les orbites sont contenues dans des droites.

Remarque 3.2.1. *Dans les trois premiers cas, les T sont conjugués topologiquement. Quand on a T_1 dans le cas i, T_2 dans le cas j, $i, j = 1, 2, i \neq j$, T_1 et T_2 sont conjugués topologiquement dans la sphère $\mathbb{S}^2$, mais pas dans $\mathbb{R}^2$. Attention : dans le cas 4, les T ne sont même pas conjugués entre eux ! Quel est l'invariant de conjugaison ? dans le cas 6, la conjugaison est linéaire.*

Les trois premiers cas sont "stables dans l'espace des paramètres" : toute application proche de T (dans $M_2(\mathbb{R})$) est topologiquement conjuguée à T. En fait ils sont même structurellement stables, au sens de la section précédente. Les autres sont instables : la moindre perturbation de T (même parmi les applications linéaires) peut changer complètement la dynamique.

Exercice 21. *Trouver un système dynamique sur $\mathbb{R}P^2$ qui n'est pas conjugué à un système linéaire.*

Ce résultat est rassurant : la dynamique est plus riche qu'on ne pourrait le croire en regardant les systèmes linéaires !

3.3 Théorème de Hartman-Grobman

Comme on l'a vu plus haut, la dynamique linéaire est assez limitée. Son intérêt est qu'elle peut souvent servir de modèle local (de même que les applications linéaires différentiables servent de modèle local aux applications différentiables). Voici un exemple, mais d'abord un peu de terminologie.

On dit qu'une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ est hyperbolique si toutes ses valeurs propres (complexes) sont de module $\neq 1$ (observez que cela correspond aux cas structurellement stables des sections précédentes).

Si f est un difféomorphisme de U dans U, où U est un voisinage de 0 dans $\mathbb{R}^n$, tel que $f(0) = 0$, on dit que 0 est point fixe hyperbolique de f si la différentielleDf_0 de f en 0 est hyperbolique.

Théorème 3.3.1. *Soit f comme ci-dessus, tel que* 0 *est point fixe hyperbolique. Alors il existe un voisinage V de* 0 *dans $\mathbb{R}^n$, $V \subset U$, tel que f et Df_0 sont conjugués topologiquement dans V.*

Le cas $n = 1$ est démontré ci-dessus ; la preuve du cas général peut faire l'objet d'un exposé.

Chapter 4

Récurrence et tous ses amis

4.1 Récurrence

Définition 4.1.1. *L'ensemble ω-limite d'un point $x \in X$ est*

$$\bigcap_{n\in\mathbb{N}} \overline{\{T^k(x)/k \geq n\}}.$$

C'est tout simplement l'ensemble des valeurs d'adhérence de l'orbite de x. Les dynamiciens cherchent des informations asymtpotiques, ils s'intéressent donc essentiellement aux ensembles ω-limite, plus qu'aux orbites elle-mêmes.

Exercice 22. *Quels sont les ensembles ω-limite possibles pour une rotation? pour le doublement ?*

Si vous avez eu du mal à répondre à la question pour le doublement, c'est que la réponse est compliquée, comme le montre l'exercice suivant :

Exercice 23. *Dans cet exercice au lieu du doublement, on considère le triplement : $\mathbb{T} \to \mathbb{T}$, $x \mapsto 3x \mod 1$. Montrer qu'il existe $x \in \mathbb{T}$ dont l'ensemble ω-limite est le Cantor triadique (l'ensemble des points dont le développement en base 3 ne contient pas le chiffre 1).*

Définition 4.1.2. *On dit qu'un point $x \in X$ est récurrent s'il appartient à son propre ensemble ω-limite.*

Ce sont essentiellement les seuls points intéressants du point de vue du dynamicien !

Exercice 24. *Si X est compact, montrer qu'il existe au moins un point récurrent. Donner un exemple de système n'ayant qu'un seul point récurrent. Montrer que l'hypothèse de compacité est nécessaire.*

Solution : Supposons qu'aucun $x \in X$ n'est récurrent. D'après l'exercice 25, pour tout $x \in X$, il existe un voisinage U_x de x dans lequel x ne revient jamais. Les U_x recouvrent X qui est compact, donc il existe $\delta > 0$ (le nombre de Lebesgue du recouvrement), tel que toute boule de rayon δ est contenue dans un U_x. Donc aucun x ne revient à distance $\leq \delta$ de lui-même. Par conséquent la suite des itérés de x ne peut pas avoir de valeur d'adhérence, ce qui est impossible puisque X est compact.

Exercice 25. *Un point $x \in X$ est non-récurrent si et seulement s'il existe un voisinage U de x, dans lequel x ne revient jamais.*

Voici quelques exemples de récurrence :

Exercice 26. *Quels sont les points récurrents d'une rotation ? du doublement ?*

4.1.1 Théorème de Poincaré

La notion de mesure invariante est souvent utile. On dit qu'une mesure borélienne de probabilité μ sur X est invariante par T si pour tout borélien $A \subset X$, on a $\mu(A) = \mu(T^{-1}(A))$. Voici, à titre d'illustration, un

Théorème 4.1.3 (Poincaré, 1890)**.** *Supposons que X est compact et que T préserve une mesure borélienne de probabilité μ sur X. Alors μ-presque tout point de X est récurrent.*

Remarques:

0-la compacité de X n'est pas vraiment nécessaire (à condition quand même que la topologie de X soit à base dénombrable) si une mesure est préservée, par contre il est essentiel que la mesure soit de masse totale finie (penser aux translations dans $\mathbb{R}$).

1-si μ est la mesure de Dirac sur un point fixe, le théorème ne dit rien du tout ! Il n'est intéressant que si le support de μ est gros, par exemple si X est le cercle et μ est la mesure de Lebesgue.

2-Notons aussi que la notion de mesure n'existait pas en 1890, Poincaré (qui avait en vue des applications à la mécanique céleste) considérait des transformations de $\mathbb{R}^n$ qui préservent le volume.

3-Rappelez-vous que Poincaré cherchait à répondre à des questions qualitatives du style "une planète peut-elle s'échapper du système solaire". Ce théorème donne une réponse assez satisfaisante pour un physicien : la probabilité pour que ça arrive est nulle. Le problème bien sûr pour appliquer ça en mécanique céleste, est qu'il y a bien une mesure préservée, mais qu'elle n'est pas de volume fini. A propos du système solaire, voir des simulations récentes sur le site web de J. Laskar, qui tendent à prouver que oui, les planètes peuvent se balader allègrement dans le sytème solaire, voire même s'échapper complètement.

Proof. (du théorème) Nous aurons besoin du

Lemme 4.1.4. *Supposons que X est compact et que T préserve une mesure borélienne de probabilité μ sur X. Soit A un borélien de X. Alors l'ensemble des points de A qui ne reviennent jamais dans A, est de mesure nulle.*

Proof. (du lemme)
Considérons, pour $n \in \mathbb{N}$,

$$A_n = \bigcup_{k \geq n} T^{-k}(A).$$

Observons que les A_n forment une suite décroissante (au sens de l'inclusion) de boréliens (rappel : l'ensemble des boréliens est stable par réunion dénombrable) de même mesure (puisque $T^{-1}(A_n) = A_{n+1}$). En particulier le complémentaire de A_{n+1} dans A_n est de mesure nulle (c'est là qu'on utilise le fait que la mesure est finie).

D'autre part l'ensemble des points de A qui ne reviennent jamais dans A, est exactement

$$\begin{aligned} \{x \in A \ : \ \forall n > 0, T^n(x) \notin A\} &= \{x \in A \ : \ \forall n > 0, x \notin T^{-n}(A)\} \\ &= A \setminus A_1 = A_0 \setminus A_1, \end{aligned}$$

qui est donc de mesure nulle. □

Revenons à la preuve du théorème. C'est ici qu'on utilise le fait que la topologie de X est à base dénombrable : il existe une famille dénombrable d'ouverts U_n, telle que tout ouvert contient un des U_n (pour X métrique compact, il suffit de prendre une suite x_n dense dans X, et de prendre la famille des boules de rayons rationnels centrées en les x_n).

Pour chaque entier n, considérons l'ensemble U'_n des points de U_n qui ne reviennent jamais dans U_n, U'_n est de mesure nulle d'après le lemme précédent, donc, par additivité dénombrable de la mesure, la réunion des U'_n est de mesure nulle.

Soit maintenant x un point non-récurrent, d'après l'exercice 25, il existe un voisinage U de x dans lequel x ne revient jamais. Soit n tel que $U_n \subset U$, on a alors $x \in U'_n$.

Par conséquent l'ensemble des points non-récurrents est contenu dans la réunion des U'_n, qui est de mesure nulle. □

4.2 Minimalité

Définition 4.2.1. *On dit que T est minimale si toute orbite est dense dans X.*

La minimalité est une forme d'indécomposabilité dynamique (comme la primalité est une forme d'indécomposabilité arithmétique, la simplicité d'un groupe, une forme d'indécomposabilité algébrique...) : elle dit qu'il n'existe pas de fermé invariant, donc pas de sous-système non-trivial.

Notez que l'existence d'une orbite périodique empêche le système d'être minimal, s'il y a plus d'une orbite bien sûr.

Exercice 27. *Montrer que les rotations d'angle irrationnel (à π) sont minimales.*

Une application marrante :

Exercice 28. *Montrer qu'il existe une infinité de puissances de* 2 *qui commencent par* 777 *(en base* 10*).*

Indication Utiliser le fait que $\log_{10}(2)$ est irrationnel. Pour info, les 5 premières sont les puissances 282, 2418, 4554, 6690, et 8826-ièmes.

Exercice 29. *Le pendule est-il minimal ?*

Indication Montrer que la fonction énergie $E(x, v) = \frac{1}{2}v^2 + \cos(x)$ est cosntante le long des orbites.

Correction La bonne question à poser est "dans un niveau d'énergie donné, la dynamique du pendule est-elle minimale ?". En général, trouver une quantité conservée (une fonction constante sur les orbites) permet de décomposer le système en sous-sytèmes.

Définition 4.2.2. *On dit que T est topologiquement transitive s'il existe un point $x \in X$ d'orbite dense.*

Cette définition ressemble à une minimalité au rabais, et c'est effectivement son principal intérêt : quand on n'arrive pas à démontrer la minimalité (par exemple dans les tentatives de preuve mathématique de l'hypothèse ergodique de Boltzmann), on tente la transitivité topologique.

Exercice 30. *Montrer que le doublement et le décalage sont topologiquement transitifs.*

Indication : pour le décalage, penser au nombre de Champernowne. Pour le doublement, se ramener au décalage.

Exercice 31. *Montrer que le petit dernier de la famille quadratique ($\lambda = 4$) est topologiquement transitif.*

Indication : se ramener au doublement en observant que $f_4(\sin^2 \theta) = \sin^2 2\theta$.

En fait on peut montrer (on le fera un peu plus loin), que f_λ devient topologiquement transitive un peu avant que λ n'atteigne 4.

Exercice 32. *X est compact. Montrer que T est topologiquement transitive si et seulement si pour tous ouverts U, V de X, il existe $n \in \mathbb{N}$ tel que $T^n(U) \cap V \neq \emptyset$.*

heuristique : tout le monde rencontre tout le monde, un jour ou l'autre.
Solution $\Longrightarrow$: soit x d'orbite dense, alors il existe $k \in \mathbb{N}$, tel que $T^k(x) \in U$, et l'orbite de $T^k(x)$ est dense (exercice : l'orbite de x est dense si et seulement si l'ensemble ω-limite de x est dense), donc il existe $p \in \mathbb{N}$, tel que $T^{k+p}(x) \in V$, de sorte que $T^p(U) \cap V \neq \emptyset$.

$\Longleftarrow$: prenons une base dénombrable $U_n, n \in \mathbb{N}$ de la topologie de X, formée d'ouverts d'adhérence compacte (pour les puristes : oui, il suffirait donc de supposer X localement compact et à base dénombrable). Définissons par récurrence une suite X_n de compacts emboîtés d'intérieur non vide tels que l'orbite de tout point de X_n rencontre $U_1, \ldots, U_n$: $X_1 = \overline{U}_1$, et si X_n est défini, il existe $k(n)$ tel que $T^{k(n)}(X_n) \cap U_{n+1} \neq \emptyset$, on pose alors $X_{n+1} = X_n \cap T^{-k(n)}(U_{n+1})$.

Point de vue de l'analyse fonctionnelle

Il est souvent utile d'adopter un point de vue dual : au lieu de considérer l'action de T sur les points de X, on considère l'action de T sur les fonctions, par l'application $f \longrightarrow f \circ T$. Un des avantages est que les fonctions forment un espace vectoriel et qu'on peut donc utiliser de l'algèbre linéaire. Par exemple :

Exercice 33. *1. Montrer que si X est compact, les seules valeurs propres possibles de l'action de T sur les fonctions continues sur X sont ± 1.*

2. Montrer que si T est topologiquement transitif sur X (et X est connexe dans le cas d'un système discret), alors 1 est la seule valeur propre de l'action de T sur les fonctions continues sur X, et c'est une valeur propre simple.

3. Montrer par un exemple que -1 peut être valeur propre pour un système discret topologiquement transitif sur un espace non connexe.

En d'autres termes, aucun espoir de décomposer le système en sous-système au moyen d'une fonction constante sur les orbites.

4.3 Mélange topologique

Définition 4.3.1. *On dit que T est topologiquement mélangeante si pour tous ouverts U, V de X, il existe $N \in \mathbb{N}$ tel que $\forall n \geq N$, $T^n(U) \cap V \neq \emptyset$.*

heuristique : le lait finit par rencontrer le café, et après on ne peut plus les séparer.

Exercice 34. *Les rotations sont-elles topologiquement mélangeantes ? et le décalage ?*

Exercice 35. *Montrer que le petit dernier de la famille quadratique ($\lambda = 4$) est topologiquement mélangeant.*

Le mélange topologique implique un certain degré de chaos, illustré par cet

Exercice 36. *Montrer que si T est une isométrie ($d(T(x), T(y)) = d(x, y)$ pour tous $x, y \in X$), alors T n'est pas topologiquement mélangeante. Connaissez-vous une isométrie minimale ?*

Solution L'identité n'étant clairement pas mélangeante, on peut supposer que T n'est pas l'identité. Soit alors $x \in X$ tel que $T(x) \neq x$, soit $d = d(x, T(x))$, soit U la boule ouverte de centre x et de rayon $d/4$, et soit V la boule de centre $T(x)$ et de rayon $d/4$. Supposons T mélangeante, et prenons $N \in \mathbb{N}$ tel que pour tout $n \geq N$, on a $T^n(U) \cap V \neq \emptyset$. Notons que, puisque T est une isométrie, $T^n(U)$ est la boule ouverte de centre $T^n(x)$ et de rayon $d/4$. Alors, par l'inégalité triangulaire, pour tout $n \geq N$, $d(T^n(x), T(x)) < d/2$, et en appliquant encore l'inégalité triangulaire, pour tout $n \geq N$, $d(T^n(x), T^{n+1}(x)) < d$, ce qui contredit $d(T^n(x), T^{n+1}(x)) = d(x, T(x))$.

Exercice 37. *Mélange topologique implique transitivité topologique. Réciproque fausse.*

Chapter 5

Expansivité

Nous avons vu que la dynamique des applications contractantes est d'une pauvreté navrante : tout le monde s'accumule sur le point fixe, et puis c'est tout. Mais qu'en est-il des applications dilatantes ? Dit comme ça, ça ne semble pas très prometteur : l'application $\mathbb{R} \to R$, $x \mapsto 2x$, est dilatante (on a, pour tous $x, y \in \mathbb{R}$, $d(T(x), T(y)) = 2d(x, y)$), et sa dynamique n'est pas beaucoup plus intéressante : un point fixe répulsif (zéro), et tout le reste part à l'infini. La situation devient plus intéressante si on impose la compacité de X. Evidemment dans ce cas la dilatation ne peut être vérifiée que localement :

Exercice 38. *Montrer qu'il n'existe pas d'application (globalement) dilatante sur un compact X.*

Solution Soient d le diamètre de X et soient $x, y \in X$ tels que $d = d(x, y)$. On aurait alors $d(T(x), T(y)) > d$, ce qui est absurde.

Définition 5.0.2. *On dit que T est localement dilatante (ou* ***expansive****) si T est continue et si pour tout $x \in X$ il existe un voisinage U_x de x dans X, et $\lambda_x > 1$, tels que pour tout $y \in U_x$, on a $d(T(x), T(y)) \geq \lambda_x d(x, y)$.*

Cette notion est différente des précédentes : la récurrence, ou la minimalité, sont des notions globales, c'est à dire qu'elles supposent de connaître le comportement du système sur un temps très long, il est très difficile en général, juste en regardant T, de dire si un sytème est minimal. L'expansivité, au contraire, est une notion locale : on peut la vérifier juste en regardant T au voisinage de chaque point de X, pas besoin de calculer des orbites sur un temps très long. Pourtant cette notion locale, sous certaines hypothèses topologiques sur l'espace X, a des conséquences globales très fortes sur la dynamique de T : par exemple, elle interdit l'existence d'attracteurs simples comme des points fixes ou des orbites périodiques attractifs.

Exercice 39. *Montrer que si X est compact, on peut prendre λ indépendant de x. On parle alors d'application λ-expansive.*

Question Si X est compact, et T est expansive, alors T peut-elle être injective ?

Réponse De façon surprenante, oui. On trouvera un exemple d'un tel X (qui ferait un excellent sujet d'exposé, soit dit en passant), appelé solénoïde de Smale, p.532 du Katok.

Dans le cas où X est une variété riemannienne compacte de dimension n, on peut montrer que le déterminant Jacobien de T est $\geq \lambda^n$, et si T est injective, on a alors que le volume total de $T(X)$ est $\geq \lambda^n$ fois le volume total de X, ce qui est impossible puisque $T(X)$ est contenue dans X.

Voici une conséquence plus intéressante : si T est expansive, on a de l'imprévisibilité (au sens du papillon) : des points arbitrairement proches se retrouvent loin au bout d'un certain temps ! Le chaos pointe le bout de son nez.

Voici un exemple d'expansivité :

Exercice 40. *Montrer que le doublement (ainsi que le triplement, etc...) est expansif.*

On notera T_k, pour $k \in \mathbb{Z}$, l'application $\mathbb{T} \to \mathbb{T}$, $x \mapsto kx \mod 1$.

Ce n'est pas par hasard que dans les exemples X est le cercle : admettre une transformation expansive n'est pas à la portée de n'importe quel espace métrique, comme le prouve cet

Exercice 41. *Montrer qu'il n'existe pas d'application expansive de* $[0, 1]$ *dans lui-même.*

Solution Supposons $T : [0, 1] \longrightarrow [0, 1]$ expansive. Soient $\lambda > 1$ et $\delta > 0$ tels que T est λ-dilatante dans tout intervalle de longueur δ. Alors l'application continue $(x, t) \mapsto f(x + t\delta) - f(x)$ ne s'annule pas, pour $1 \geq t > 0$, donc elle garde un signe constant, celui de $f(\delta) - f(0)$. Supposons $f(\delta) - f(0) > 0$, l'autre cas est identique. On a alors $f(x+t\delta)-f(x) \geq \lambda t\delta$ pour tous (x, t) avec $1 \geq t > 0$. Soient maintenant $x > y$ et soit $n = \lfloor x - y\rfloor/\delta$, on a

$$\begin{aligned} f(x) - f(y) =& f(x) - f(x - \delta) + f(x - \delta) - f(x - 2\delta) + \ldots + \\ & f(x - (n - 1)\delta) - f(x - n\delta) + f(x - n\delta) - f(y) \\ & \geq n\lambda\delta + \lambda(x - n\delta - y) = \lambda(x - y), \end{aligned}$$

c'est à dire que T est dilatante, ce qui est impossible par l'exercice 38.

Plus généralement, une application expansive d'une variété compacte dans elle-même est un revêtement, et admettre un auto-revêtement est une propriété spéciale ; parmi les surfaces compactes orientables, seul le tore en admet.

Question : lien avec la transitivité topologique ? le mélange topologique ?

Sous cette forme, la question est posée de façon un peu trop générale :

Exercice 42. *Construire un système (X, T) avec X compact, T expansive, qui n'est pas topologiquement transitif.*

Solution : X est un bouquet de deux cercles qui se touchent en l'élément neutre, T est le doublement sur chacun des cercles.

En revanche le théorème 5.2.1 donne une réponse positive dans le cas où X est le cercle. Une question intéressante (mais qui dépasse largement le cadre de ce cours) : quels sont les X sur lesquels la réponse est positive (i.e. expansif implique topologiquement transitif) ?

Exercice 43. *Montrer que toute application expansive du cercle a un point fixe, et même des orbites périodiques de toutes périodes.*

Indication : relever à $\mathbb{R}$ et utiliser le théorème des valeurs intermédiaires.
Avant d'approfondir la question, un petit détour topologique.

5.1 Degré d'une transformation du cercle

Notons π la projection canonique $\mathbb{R} \longrightarrow \mathbb{T} = \mathbb{R}/\mathbb{Z}$. Je noterai f au lieu de T les applications du cercle, afin de réserver la majuscule aux relèvements : rappelez-vous (que si f est une application continue de $\mathbb{T}$, on appelle relèvement de f à $\mathbb{R}$, une application $F : \mathbb{R} \longrightarrow \mathbb{R}$ telle que $\forall x \in \mathbb{R}, \pi(F(x)) = f(\pi(x))$. Alors, pour tout $x \in \mathbb{R}$, il existe $n(x) \in \mathbb{Z}$ tel que $F(x+1) = F(x) + n(x)$. Par continuité de F, l'application $x \mapsto n(x)$ est localement constante, et par connexité de $\mathbb{R}$ elle est constante.

Définition 5.1.1. *On appelle degré de f l'entier n tel que $F(x+1) = F(x) + n$ pour tout $x \in \mathbb{R}$.*

Evidemment, cette définition n'a de sens qu'avec cet

Exercice 44. *Montrer que le degré ne dépend pas du relèvement !*

Exercice 45. *Montrer que toute application du cercle, de degré non nul, est surjective.*

Exercice 46. *Monter que si F relève f, et G relève g, alors $F \circ G$ relève $f \circ g$.*

Exercice 47. *Montrer que le degré de $f \circ g$ est $deg(f).deg(g)$.*

Exercice 48. *Montrer que le degré est invariant par conjugaison topologique.*

Exercice 49.
1. *Montrer que tout homéomorphisme du cercle est de degré ± 1.*
2. *Construire une application du cercle dans lui-même, de degré 1, qui n'est pas un homéomorphisme.*

3. *Montrer qu'une application du cercle dans lui-même, de degré 1, admettant un relèvement strictement monotone (forcément croissant), est un homéomorphisme.*

Exercice 50. *Montrer que si deux transformations du cercle sont suffisamment proches, dans la topologie uniforme, alors elles ont même degré.*

Exercice 51. *Montrer qu'une application h de degré 1 qui commute avec T_k est l'identité.*

Solution Relevons h à $\mathbb{R}$ par une application H. Observons tout d'abord que $H(0) = 0$ puisque $H(0) = kH(0)$. On va montrer que H est l'identité sur les rationnels k-adiques, et conclure par continuité de H. Soient $p, n \in \mathbb{N}$, et calculons $H(p/k^n)$. Puisque H commute avec T_k, on a

$$p = H(0) + p = H(p) = k^n H(p/k^n).$$

Revenons aux applications expansives.

5.2 Stabilité structurelle des applications expansives

Théorème 5.2.1. *Toute application de $\mathbb{T}$, expansive et de degré k, est topologiquement conjuguée à T_k.*

Notons que ce résultat est plus fort que la simple stabilité structurelle puisqu'on ne suppose pas T différentiable, mais seulement continue. De plus il n'y a aucune hypothèse de proximité, ni C^1 ni même C^0, avec T_k.

Avant de démontrer le théorème, voici une proposition plus simple mais qui contient l'idée essentielle de la preuve (vous trouverez deux preuves complètes p. 73 et suivantes du Katok) :

Proposition 5.2.2. *Toute application de $\mathbb{T}$, de degré k, est topologiquement semi-conjuguée à T_k, par une application de degré* 1.

Pour comprendre pourquoi la conclusion de la proposition est plus faible que celle du théorème, penser à l'exercice 49. D'un autre côté la proposition ne suppose pas l'expansivité.

Proof. (de la Proposition 5.2.2) Soit $f : \mathbb{T} \to \mathbb{T}$, et de degré k. Il s'agit de résoudre l'équation fonctionnelle

$$T_k \circ h = h \circ f \tag{5.1}$$

avec h de degré 1. Soit $\mathcal{C}$ l'ensemble des applications continues H de $\mathbb{R}$ dans lui-même, telles que $H(x + 1) = H(x) + 1$. C'est exactement l'ensemble des relèvements à $\mathbb{R}$ des applications de degré 1 du cercle dans lui-même. Munissons $\mathcal{C}$ de la distance uniforme :

$$d(H_1, H_2) = \max_{x \in \mathbb{R}} |H_1(x) - H_2(x)|.$$

Exercice 52. *Montrer que cette distance est bien définie, et que $\mathcal{C}$, muni de la distance uniforme, est complet.*

Choisissons un relèvement F de f à $\mathbb{R}$. L'équation (5.1) revient alors à

$$\forall x \in \mathbb{R}, kH(x) = H(F(x)),$$

avec $H \in \mathcal{C}$. Si vous ne voyez pas pourquoi cela revient à (5.1), pensez à l'exercice 46.

Considérons l'application

$$\begin{aligned} \mathcal{F} : \mathcal{C} &\longrightarrow \mathcal{C} \\ H &\longmapsto \frac{1}{k} H \circ F. \end{aligned}$$

Ce que nous cherchons, c'est donc un point fixe de $\mathcal{F}$. Maintenant, observons que $\mathcal{F}$ est contractante, d'un facteur $1/k$, pour la distance uniforme :

$$(\forall x \in \mathbb{R}, |H_1(x) - H_2(x)| \leq \epsilon) \Longrightarrow \left(\forall x \in \mathbb{R}, \frac{1}{k}|H_1(F(x)) - H_2(F(x))| \leq \frac{\epsilon}{k}\right).$$

On conclut alors avec l'exercice 6. ☐

Proof. (du théorème 5.2.1) A présent il nous reste à démontrer que le h donné par la Proposition 5.2.2 est un homéomorphisme quand T est expansive. Soit $f : \mathbb{T} \to \mathbb{T}$, expansive et de degré k. Conservons les notations de la preuve de 5.2.2. L'idée est de résoudre l'équation fonctionnelle associée à h^{-1} : on considère l'application

$$\begin{aligned} \mathcal{G} : \mathcal{C} &\longrightarrow \mathcal{C} \\ H &\longmapsto F^{-1} \circ H \circ T_k \end{aligned}$$

et on cherche un point fixe de $\mathcal{G}$.

Soient $\lambda > 1, \delta > 0$ tels que

$$\begin{aligned} \forall x, y \in \mathbb{R}, \delta > x - y > 0 &\Longrightarrow F(x) - F(y) \geq \lambda(x - y) \\ \text{et } x - y \geq \delta &\Longrightarrow F(x) - F(y) \geq \lambda\delta. \end{aligned}$$

Il est facile de voir qu'en fait $\forall x, y \in \mathbb{R}, F(x) - F(y) \geq \lambda(x - y)$, de sorte que F^{-1} est contractante, d'un facteur λ^{-1}, donc $\mathcal{G}$ est également contractante, d'un facteur λ^{-1}. On obtient donc un point fixe H_g de $\mathcal{G}$. En notant H_f le point fixe de $\mathcal{F}$, on a alors

$$F \circ H_g = H_g \circ T_k \text{ et } H_f \circ F = T_k \circ H_f$$

donc

$$H_f \circ H_g \circ T_k = H_f \circ F \circ H_g = T_k \circ H_f \circ H_g$$

et on conclut alors avec l'exercice 51. ☐

Applications du théorème 5.2.1

Exercice 53. *Montrer que toute application expansive du cercle est topologiquement transitive.*

Exercice 54. *Montrer que toute application expansive du cercle est topologiquement mélangeante.*

Exercice 55. *Montrer que l'expansivité est une condition ouverte dans l'espace des applications du cercle dans lui-même, de degré k, pour la topologie C^1 (pas C^0 !).*

En déduire le

Théorème 5.2.3. *Les transformations expansives du cercle sont structurellement stables.*

Ce théorème est une illustration de plus du principe philosophique suivant : plus une application est chaotique, plus elle est structurellement stable. Il n'y a pas plus robuste que le chaos ! Attention quand même, c'est un principe heuristique, pas un théorème, faute de définition précise du chaos.

Exercice 56. *Montrer que le chaos dans la famille quadratique arrive strictement avant $\lambda = 4$.*

Chapter 6

Ergodicité

6.1 Définition

La théorie de la mesure va nous permettre de quantifier (au sens de rendre quantitatives) les notions déjà vues. C'est bien joli de savoir qu'une orbite est récurrente, mais avec quelle fréquence revient-elle ? (le verbe associé à l'adjectif récurrent est revenir. Le verbe récurrer n'existe pas. Just sayin') Si une comète revient tous les 10 ans, ce n'est pas pareil que si elle revient tous les 1000 ans, et ce n'est pas non plus pareil que si elle revient la première fois au bout de 10 ans, la deuxième au bout de 100 ans, la troisième au bout de 1000 ans... la n-ième au bout de 10^n ans. C'est la fréquence asymptotique qu'il s'agit de définir. Ici, la notion de mesure invariante est fondamentale. On va commencer par quelques rappels :

-le théorème de représentation de Riesz qui dit que si X est un espace métrique compact, alors l'ensemble $M(X)$ des mesures boréliennes de probabilité X s'identifie à l'ensemble des formes linéaires sur $C(X)$ (l'ensemble des fonctions continues sur X), qui sont positives sur les fonctions positives, et qui valent 1 sur la fonction constante 1.

-cet espace peut-être muni d'une topologie au nom barbare (topologie faible-*) et à la définition très simple : c'est tout simplement la topologie de la convergence simple, on dit qu'une suite de mesures μ_n converge vers μ si pour toute fonction $f \in C(X)$, on a $\int f d\mu_n \longrightarrow \int f d\mu$.

-le théorème de Banach-Alaoglu dit alors que $M(X)$, muni de la topologie faible-*, est compact.

-mais ce n'est pas tout : $C(X)$ est un espace vectoriel, donc on peut parler de convexité d'une partie de $C(X)$, et justement, $M(X)$ est convexe : décidément, c'est quelqu'un de tout à fait fréquentable.

Exercice 57. *Si X est compact, démontrer l'existence de mesures invariantes.*

Indication Utiliser le théorème de Banach-Alaoglu.

On notera désormais $M_T(X)$ l'ensemble des mesures boréliennes de probabilité T-invariantes sur X.

Exercice 58. *Le support d'une mesure invariante est alors une partie T-invariante de X.*

Définition 6.1.1. *Lorsqu'il n'existe pas de parties T-invariantes de X de mesure différente de 0 et de 1, on dit que la mesure μ est **ergodique** par rapport à T, ou, de façon équivalente, que T est ergodique pour μ.*

L'ergodicité peut-être vue comme une condition d'indécomposabilité, au sens de la mesure (pas de borélien invariant non trivial), de même que la minimalité est une condition d'indécomposabilité topologique (pas de fermé invariant non trivial).

Exercice 59. *Si X est compact, démontrer l'existence de mesures ergodiques.*

Solution : $M_T(X)$ est un convexe compact pour la topologie faible*. Nous allons montrer que les mesures ergodiques sont exactement les points extrémaux de ce convexe, puis utiliser le théorème de Carathéodory qui dit qu'un convexe compact possède des points extrémaux.

Montrons d'abord qu'un point extrémal est une mesure ergodique. Soit μ un point extrémal de $M_T(X)$, et soit B un borélien T-invariant, il s'agit de montrer que $\mu(B) = 0$ ou 1. Supposons que $\mu(B) \neq 0, 1$, on peut alors considérer les mesures conditionnelles à B et $X \setminus B$,

$$\mu_B(.) = \frac{\mu(. \cap B)}{\mu(B)} \text{ et } \mu_{X\setminus B}(.) = \frac{\mu(. \cap X \setminus B)}{\mu(X \setminus B)}.$$

On a alors $\mu = \mu(B)\mu_B + \mu(X \setminus B)\mu_{X\setminus B}$, ce qui contredit l'extrémalité de μ, donc $\mu(B) = 0$ ou 1, donc μ est T-ergodique.

Ceci suffit pour démontrer l'existence de mesures ergodiques, mais pour votre culture il est bon de savoir également qu'une mesure ergodique est un point extrémal. Soit μ T-ergodique, et supposons que μ est une combinaison convexe non-triviale, $\mu = t\lambda + (1-t)\nu$, avec λ et ν des éléments distincts de $M_T(X)$. Puisque $\lambda \neq \nu$, il existe un borélien B tel que $\lambda(B) \neq \nu(B)$. Si B est invariant, c'est fini, puisqu'on a alors $\mu(B) \neq 0, 1$, contredisant l'ergodicité. Si B n'est pas invariant, on considère l'ensemble des points de B qui reviennent une infinité de fois dans B, c'est à dire

$$B_\infty = \bigcap_{n\in\mathbb{N}} T^{-n}(B),$$

B_∞ est clairement invariant, et d'après le théorème 4.1.3, B_∞ est de mesure pleine dans B, ce qui permet de conclure en utilisant B_∞ à la place de B.

Les supports de mesures ergodiques sont des ensembles invariants particulièrement intéressants.

Exercice 60. *Montrer que les points fixes et orbites périodiques sont supports de mesures ergodiques.*

Indication Utiliser le théorème de Radon-Riesz qui dit qu'une mesure sur X n'est rien d'autre qu'une forme linéaire positive sur l'espace des fonctions continues sur X (muni de la topologie uniforme).

Définition 6.1.2. *On que dit que T est uniquement ergodique s'il existe une seule mesure borélienne de probabilité T-invariante sur X.*

cette terminologie est justifiée par l'

Exercice 61. *Montrer que si T est uniquement ergodique, la mesure T-invariante est ergodique par rapport à T.*

Solution : si on a une partie invariante $A \subset X$ de mesure différente de 0 et 1, la mesure "μ sachant A" ($\mu_A(B) = \mu(B \cap X)/\mu(A)$) est invariante, et différente de μ, ce qui contredit l'unique ergodicité.

Exercice 62. *Montrer que si T est uniquement ergodique, et μ est la mesure de probabilité invariante, alors T, restreinte au support de μ, est minimale.*

Solution : l'adhérence de n'importe quelle orbite supporte une mesure invariante, donc est égale au support de μ.

Exercice 63. *L'unique ergodicité implique-t-elle la minimalité ? et si T est un homéomorphisme de X ?*

Réponse : sous cette forme, la question est mal posée. Une application contractante sur un compact est uniquement ergodique, et la mesure ergodique est le Dirac sur le point fixe. La bonne question est "l'unique ergodicité implique-t-elle la minimalité de la dynamique restreinte au support de la mesure invariante ?" La réponse est oui, voici pourquoi.

Soit x dans le support de la mesure ergodique μ, et montrons que le support de μ tout entier est contenu dans l'ensemble omega-limite $\omega(x)$ de x.

Exercice 64. *Montrer que l'unique ergodicité est invariante par conjugaison topologique.*

Exercice 65. *Montrer que si T est uniquement ergodique, alors 1 est valeur propre simple de l'action duale sur les fonctions intégrables.*

6.2 Ergodicité des rotations

Exercice 66. *Montrer que la mesure de Lebesgue sur le cercle est invariante par les rotations, et par le doublement.*

Théorème 6.2.1. *La rotation R_α d'angle α est uniquement ergodique si et seulement si $\alpha/\pi \notin \mathbb{Q}$, et dans ce cas, la mesure ergodique est la mesure de Lebesgue. Si $\alpha/\pi \in \mathbb{Q}$, la mesure de Lebesgue n'est pas ergodique.*

Proof. Débarrassons-nous d'abord du sens facile : si α/π est rationnel, en prenant un petit voisinage d'une orbite périodique, on obtient un ensemble invariant de mesure ni nulle, ni pleine.

A présent supposons α/π irrationnel.

Nous allons utiliser la caractérisation suivante de la mesure de Lebesgue:

Proposition 6.2.2. *La mesure de Lebesgue est la seule mesure borélienne sur $\mathbb{R}$ qui est invariante par translation et donne une masse* 1 *à l'intervalle* $[0, 1]$.

Si vous connaissez les groupes de Lie, la mesure de Lebesgue est la mesure de Haar du groupe $\mathbb{R}/\mathbb{Z}$. Voici une preuve élémentaire (les détails sont à faire en exercice).

1-il suffit de montrer que $\mu([a, b]) = |b - a|$ pour tous $a, b \in \mathbb{R}$, puisque les intervalles engendrent les ouverts, donc la tribu des boréliens.

2-l'invariance par translation entraîne que c'est vrai pour tout intervalle dont les bornes sont rationnelles :

$$\mu\left(\left[\frac{p}{q}, \frac{r}{s}\right]\right) = \mu\left(\left[0, \frac{r}{s} - \frac{p}{q}\right]\right) = \mu\left(\left[0, \frac{P}{Q}\right]\right) = P\mu\left(\left[0, \frac{1}{Q}\right]\right) = \frac{P}{Q}$$

où on a posé $P = ps - rq$, $Q = qs$, et utilisé le fait que la mesure de P (ou Q) intervalles identiques juxtaposés est égale à P (ou Q) fois la mesure d'un des intervalles.

3-on conclut ensuite en approximant a, b par des rationnels.

A présent démontrons le théorème. Supposons une mesure de probabilité sur le cercle invariante par une rotation irrationnelle R_α ; en la relevant à $\mathbb{R}$ on obtient une mesure sur $\mathbb{R}$, invariante par le groupe (additif) $\mathbb{Z} + \alpha\mathbb{Z}$, et qui donne une masse 1 à l'intervalle $[0, 1]$. En utilisant le fait que le groupe $\mathbb{Z} + \alpha\mathbb{Z}$ est dense dans $\mathbb{R}$, on obtient que la mesure est en fait invariante par toutes les translations. □

6.3 Théorème de Birkhoff

La notion d'ergodicité peut être vue comme une notion faible de chaos, c'était le point de vue originel de Boltzmann à qui on doit le mot : les orbites sont tellement compliquées que la moyenne spatiale d'une fonction, est égale à sa moyenne sur une orbite (en maths, c'est le théorème ergodique de Birkhoff, ci-dessous). Toutefois l'exemple des rotations montre que l'ergodicité n'implique aucune sorte d'imprévisibilité, au sens du papillon.

Théorème 6.3.1 (G. Birkhoff, 1931)**.** *Supposons que T préserve la mesure borélienne de probabilité μ, et soit f une fonction dans $L^1(X, \mu)$. Alors la moyenne (temporelle) de f sur l'orbite de x :*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(T^k(x))$$

existe pour presque tout $x \in X$.

Notons que μ n'est pas supposée T-ergodique dans l'énoncé du théorème. Si c'est dans le cas, avec l'exercice 65 on déduit immédiatement la forme classique du théorème (celle qui est utilisée par les physiciens) : "moyenne temporelle égale moyenne spatiale".

En particulier, en applicant le théorème à la fonction caractéristique d'une partie mesurable E de X, on obtient que la fréquence asymptotique des visites dans E de presque tout point de X, est égale à la mesure de E. En d'autres termes, presque toute orbite est équidistribuée : le temps qu'elle passe dans E est proportionnel à la place que prend E (relativement à la mesure μ).

Nous ne démontrerons pas le théorème 6.3.1 (**exposé** possible), mais la version L^2, due à Von Neumann, beaucoup plus simple et légèrement plus faible puisque la convergence L^2 n'entraîne pas la convergence presque partout, mais seulement la convergence presque partout d'une sous-suite (exemple ci-dessous).

Exemple pour votre culture: $X = [0, 1]$, f_n est une vague qui se promène de gauche à doite et de droite à gauche en rapetissant, pas en hauteur, mais en largeur.

6.4 Théorème de Von Neumann

Rappelons que l'espace de Hilbert $L^2(X, \mu)$ (fonctions de carré intégrable sur X) est muni d'un produit scalaire

$$\forall f, g \in L^2(X, \mu), \langle f, g\rangle = \int_X f(x)g(x)d\mu(x),$$

et donc d'une norme

$$N(f) = \sqrt{\int_X f(x)^2 d\mu(x)}.$$

On dit qu'une suite f_n dans $L^2(X, \mu)$ converge vers $f \in L^2(X, \mu)$ si $N(f_n - f) \to 0$. On note

$$\begin{aligned} U : L^2(X, \mu) &\longrightarrow L^2(X, \mu) \\ f &\longmapsto f \circ T. \end{aligned}$$

Exercice 67. *Supposons que T préserve la mesure μ. Montrer que U est une isométrie (i.e. préserve la norme) de $L^2(X,\mu)$. Montrer que si T est inversible, alors U aussi est inversible, et calculer son inverse.*

Vous allez me demander, mais comment U pourrait-il ne pas être inversible, puisque c'est une isométrie ? C'est la magie de l'hôtel de Hilbert !

Exercice 68. *Montrer que le sous-espace des fonctions constantes est contenu dans Ker$(U-I)$.*

On montrera la réciproque comme corollaire du théorème ergodique de Von Neumann.

Nous aurons besoin de ce petit fait sur les espaces de Hilbert (espaces vectoriels munis d'un produit scalaire et complets pour la norme induite par le produit scalaire, les exemples-types sont $L^2(X,\mu)$, ou l'espace des suites réelles de carré sommable):

Exercice 69. *Si H est un espace de Hilbert, et si F est un sous-espace de H, alors $H = \overline{F} \oplus F^\perp$, où $\overline{F}$ désigne l'adhérence de F.*

Exercice 70. *Montrer que le sous-espace Ker$(U-I)$ est fermé dans $L^2(X,\mu)$, et en déduire que la projection orthogonale (s'il y a produit scalaire, il y a projection orthogonale !) sur Ker$(U-I)$ est bien définie.*

Dans la suite on notera P la projection orthogonale sur $\mathrm{Ker}(U-I)$.

Exercice 71. *Montrer que, pour $f \in L^2(X,\mu)$, f est constante sur les orbites de T si et seulement si $f \in$ Ker$(U-I)$, c'est à dire $P(f) = f$.*

Théorème 6.4.1 (J. Von Neumann, 1932)**.** *Supposons que T préserve la mesure μ, et soit f une fonction dans $L^2(X,\mu)$. Alors on a, dans $L^2(X,\mu)$, la limite suivante :*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} U^k(f) = P(f).$$

Notons que l'usage m'oblige à donner les dates de *publication*, mais le théorème de Von Neumann est en réalité plus ancien-de quelques mois-que celui de Birkhoff.

6.5 Applications du théorème 6.4.1

Notons aussi que, pas davantage que dans le théorème de Birkhoff, le mot "ergodique" n'apparaît dans le théorème de Von Neumann. Maintenant voyons ce qu'il se passe quand T est ergodique

Proposition 6.5.1. *T ergodique $\iff \forall f \in L^2(X,\mu)$, $P(f)$ est constante μ-presque partout (nécessairement égale à $\int f d\mu$).*

Proof. Supposons T ergodique, et soit $f \in L^2(X, \mu)$. La fonction $P(f)$ est T-invariante, donc ses ensembles de niveau aussi, donc pour tout $t \in \mathbb{R}$, $\{x \in X : f(x) \geq t\}$ est invariant. Par conséquent la fonction $t \mapsto \mu(\{x \in X : f(x) \geq t\})$ est une fonction décroissante qui ne prend que les valeurs 0 et 1, donc il existe $t_0 \in \mathbb{R}$ tel que $\mu(\{x \in X : f(x) \geq t\}) = 0$ pour $t > t_0$, et $\mu(\{x \in X : f(x) \geq t\}) = 1$ pour $t < t_0$. Donc $P(f)(x) = t_0$ pour μ-presque tout x, et on voit que $t_0 = \int f d\mu$ en intégrant $P(f)$.

Réciproquement, supposons que $\forall f \in L^2(X, \mu)$, $P(f)$ est constante μ-presque partout, et soit A un borélien de X, invariant par T. Soit f la fonction caractéristique de A, alors f est invariante par T, donc $P(f) = f$, donc f est constante μ-presque partout, ce qui signifie $\mu(A) = 0$ ou 1. □

6.5.1 Ergodicité du doublement

Théorème 6.5.2. *Le doublement est ergodique pour la mesure de Lebesgue.*

Proof. On va utiliser le critère d'ergodicité 6.5.1 : il s'agit de montrer qu'une fonction L^2 invariante par le doublement T_2 est constante Lebesgue-presque partout. Soit donc une fonction f, L^2 et T_2-invariante, et décomposons-la en série de Fourier :

$$f(x) = \sum_{k \in \mathbb{Z}} a_k \exp(2\pi k x),$$

de sorte que

$$f(T_2(x)) = \sum_{k \in \mathbb{Z}} a_k \exp(4\pi k x),$$

ce qui implique l'égalité des coordonnées dans la base hilbertienne des fonctions $x \mapsto \exp(2\pi k x)$:

$$\forall k \in \mathbb{Z}, a_k = a_{2k}. \tag{6.1}$$

Mais puisque f est L^2, la suite a_k est de carré sommable, donc 6.1 entraîne $a_k = 0$ pour tout $k \neq 0$, c'est à dire que f est constante Lebesgue-presque partout. □

6.6 Démonstration du théorème 6.4.1

La raison pour laquelle la preuve du théorème de Von Neumann est plus simple que celle du théorème de Birkhoff est que nous allons pouvoir utiliser la merveilleuse théorie des espaces de Hilbert, dans lesquels (presque) tout se passe comme en dimension finie.

Rappelons tout d'abord la notion d'adjoint dans un espace hilbertien : si H est un espace de Hilbert, et si f est un endomorphisme continu de H, alors l'adjoint f^* de f est l'unique endomorphisme de H tel que

$$\forall x, y \in H, \langle x, f(y) \rangle = \langle f^*(x), y \rangle.$$

Exercice 72. *Montrer que f^* est bien défini pour tout f, et que $f^* = f^{-1}$ si f est une isométrie* **et** *un automorphisme.*

Exercice 73. *Montrer que l'application $f \mapsto f^*$ est linéaire et involutive. Montrer que pour tout f, $\ker f^* = \overline{Imf}^{\perp}$.*

Exercice 74. *Montrer que si f est une isométrie de H (pas nécessairement surjective), pour tout $x \in H$, $f(x) = x$ équivaut à $\langle x, f(x)\rangle = N(x)^2$.*

Indication : cas d'égalité dans Cauchy-Schwarz.

A partir de maintenant on revient à l'espace de Hilbert $L^2(X, \mu)$, muni de l'isométrie U.

Exercice 75. *Montrer que $\ker(U - I)^* = \ker(U - I)$. En déduire que $\ker(U - I) = Im(U - I)^{\perp}$, puis que*

$$L^2(X, \mu) = \ker(U - I) \oplus \overline{Im(U - I)}.$$

A présent nous sommes en mesure de démontrer le théorème. Soit $f \in L^2(X, \mu)$, écrivons $f = P(f) + g$, où $g \in \overline{\mathrm{Im}(U - I)}$. On a, pour tout n,

$$\frac{1}{n}\sum_{k=1}^{n} U^k(f) = \frac{1}{n}\sum_{k=1}^{n} U^k(P(f)) + \frac{1}{n}\sum_{k=1}^{n} U^k(g),$$

et puisque $P(f) \in \ker(U - I)$, on a

$$\frac{1}{n}\sum_{k=1}^{n} U^k(P(f)) = P(f).$$

Fixons $\epsilon > 0$ et prenons un élément g_ϵ de $\mathrm{Im}(U - I)$ tel que $N(g - g_\epsilon) < \epsilon$. Observons que, puisque U est une isométrie, on a, en utilisant l'inégalité triangulaire, pour tout n,

$$N\left(\frac{1}{n}\sum_{k=1}^{n} U^k(g_\epsilon) - \frac{1}{n}\sum_{k=1}^{n} U^k(g)\right) \leq \epsilon.$$

Puisque $g_\epsilon \in \mathrm{Im}(U - I)$, il existe h_ϵ tel que $g_\epsilon = U(h_\epsilon) - h_\epsilon$, donc

$$\frac{1}{n}\sum_{k=1}^{n} U^k(g_\epsilon) = \frac{1}{n}\left(U^{n+1}(h_\epsilon) - h_\epsilon\right).$$

et il existe n_ϵ tel que $\forall n \geq n_\epsilon$,

$$N\left(\frac{1}{n}\sum_{k=1}^{n} U^k(g_\epsilon)\right) < \epsilon.$$

Alors, $\forall n \geq n_\epsilon$,

$$\begin{aligned} N\left(\frac{1}{n}\sum_{k=1}^{n} U^k(f) - P(f)\right) &= N\left(\frac{1}{n}\sum_{k=1}^{n} U^k(g)\right) \\ &\leq N\left(\frac{1}{n}\sum_{k=1}^{n} U^k(g_\epsilon) - \frac{1}{n}\sum_{k=1}^{n} U^k(g)\right) + N\left(\frac{1}{n}\sum_{k=1}^{n} U^k(g_\epsilon)\right) \\ &< \epsilon + \epsilon. \end{aligned}$$

Puisqu'ϵ est arbitraire, ceci achève la preuve du théorème. □

6.7 Miscellanées

6.7.1 Application du théorème 6.3.1

Proposition 6.7.1. *Si T est ergodique pour μ, alors T, restreinte au support de μ, est topologiquement transitive.*

Proof. Soit $(U_i)_{i\in\mathbb{N}}$ une base dénombrable de la topologie du compact métrique $\mathrm{supp}(\mu)$. Par définition du support on a $\mu(U_i) > 0 \forall i$. Soit f_i la fonction caractéristique de U_i, d'après le théorème 6.3.1, il existe $R_i \subset X$ de mesure pleine tel que $\forall x \in R_i$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} f_i(T^k(x)) = \mu(U_i)$$

en particulier l'orbite de tout $x \in R_i$ visite U_i une infinité de fois. Prenons $R = \cap_{i\in\mathbb{N}} R_i$, alors R est de mesure pleine et tout $x \in R$ visite tous les U_i, donc l'orbite de tout point de R est dense dans $\mathrm{supp}(\mu)$. □

Loi forte des grands nombres

L'énoncé qui suit est simplifié pour que la démonstration soit immédiate, en fait il n'y a pas besoin de supposer que les variables aléatoires suivent une loi de Bernoulli.

Théorème 6.7.2 (Borel, 1909)**.** *Soit X_n une suite de variables aléatoires indépendantes sur un espace probabilité $(\Omega, \mathcal{A}, \mu)$, suivant toutes la même loi de Bernoulli de paramètre p (pensez-y comme une suite de tirages à pile ou face avec la même pièce). Alors presque sûrement la moyenne empirique tend vers p.*

Proof. On considère chaque réalisation de la suite X_n comme une suite de 0 et de 1. On munit l'ensemble $\{0,1\}^{\mathbb{N}}$ de la mesure-produit. Le décalage est ergodique pour cette mesure (exercice ; indication : il est conjugué, en dehors d'une partie de mesure nulle, au doublement sur le cercle). La fonction "moyenne empirique" est définie presque partout, mesurable (comme limite simple de fonctions mesurables) et bornée sur $\{0,1\}^{\mathbb{N}}$ qui est compact pour la topologie produit, donc la moyenne empirique est intégrable. La moyenne, pour la mesure-produit, de la fonction "moyenne empirique", est p (exercice; indication : Fubini). Donc l'ensemble des suites de moyenne empirique p est de mesure pleine. □

Question : que dit le théorème de Von Neumann dans cette situation?

Exercice 76. *Montrer que presque tout nombre réel est un nombre-univers, comme le nombre de Champernowne, au sens où son développement décimal (ou binaire, ou ce que vous voulez) contient toutes les suites finies possibles.*

Exercice 77. *Montrer qu'en moyenne, à peu près une puissance de* 2 *sur* 2000 *commence par* 777.

Exposé pour quelqu'un de motivé : preuve de l'ergodicité (pour quelle mesure ?) du flot géodésique (et horocyclique ?) sur les surfaces hyperboliques (fermées).

6.8 Mélange

Nous avons vu que l'ergodicité est le pendant mesuré de la transitivité topologique, et que l'unique ergodicité est le pendant mesuré de la minimalité. Voici maintenant le pendant mesuré du mélange topologique.

Dans cette section on suppose que T préserve une mesure borélienne de probabilité μ sur X.

Définition 6.8.1. *On dit que T est mélangeante relativement à la mesure μ, si pour toutes parties mesurables A et B de X, on a*

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B).$$

En termes probabilistes, cela signifie que pour tout $x \in X$, les évènements $T^n(x) \in A$ et $x \in B$ sont "asymptotiquement indépendants". En ce sens le mélange est la meilleure notion de chaos (probabiliste): une information initiale ne donne aucune information asymptotique.

Attention quand même : cette notion n'est pas très intéressante si le support de μ est tout petit ! Par exemple, si T admet un point fixe, T est mélangeante par rapport à la mesure de Dirac supportée par le point fixe, mais ça ne nous dit pas grand-chose sur la dynamique de T.

Il n'est pas clair pour autant que le mélange est la meilleure notion de chaos : par exemple, le mélange n'implique pas la positivité de l'entropie. Un contre-exemple très intéressant est donné par le flot horocyclique des surfaces hyperboliques, ça peut faire un excellent exposé, ou, pour un étudiant ambitieux, la deuxième partie d'un exposé sur l'ergodicité du flot géodésique des surfaces hyperboliques.

Le consensus parmi les experts est qu'il n'y a pas vraiment de meilleure notion de chaos, il y a plusieurs notions (mélange, entropie, et en général on y rajoute la densité des points périodiques) qui jouent chacune un rôle dans la formation du chaos. En gros, c'est un peu comme la vraie vie (par opposition aux maths).

En général le décalage (ou le doublement) est considéré comme le prototype du système chaotique (il coche toutes les cases ; en tous cas, à ma connaissance on ne fait pas plus chaotique) mais beaucoup d'autres systèmes méritent aussi d'être appelés chaotiques sans pour autant "contenir" (donner un sens à ce mot...) un décalage.

Exercice 78. *Si T est mélangeante relativement à la mesure μ, alors le support de μ est invariant par T, et la restriction de T au support de μ est topologiquement mélangeante.*

Plus difficile : construire une transformation topologiquement mélangeante qui n'est mélangeante par rapport à aucune mesure de support total.

Exercice 79. *Montrer que le décalage est mélangeant relativement à la mesure naturelle.*

Exercice 80. *Montrer que le mélange entraîne l'ergodicité.*

lien avec l'entropie ? avec l'ergodicité ?

Chapter 7

Homéomorphismes du cercle

Dans ce chapitre on va arrêter (momentanément) de définir de nouveaux concepts, et étudier un des grands succès initiaux de la théorie des systèmes dynamiques : la classification, par Poincaré (ca. 1905), des homéomorphismes du cercle, à conjugaison topologique près.

Pour se faire une idée de l'exploit, il faut voir que le groupe Homeo($\mathbb{T}$) des homéomorphismes du cercle est un groupe énorme, de dimension infinie, et même indénombrable (en gros, il est localement pareil que l'espace des fonctions continues).

Le génie de Poincaré a été de réaliser que cet énorme machin est tout mou, et le squelette n'est autre que le groupe des rotations, c'est à dire $\mathbb{T}$ lui-même, qui est de dimension 1, et très bien compris. Bon là je mens un petit peu. Mais c'est l'idée.

Une autre raison pour laquelle ce résultat (la classification de Poincaré) est très fort, est que nous n'avons aucune hypothèse locale de type hyperbolicité, ou expansivité. Notre seule arme sera la préservation de l'ordre.

7.1 Nombre de rotation

Pour des raisons pratiques on préfère travailler avec des applications de $\mathbb{R}$ dans $\mathbb{R}$ qu'avec des homéos de $\mathbb{T}$. Notons π la projection canonique $\mathbb{R} \longrightarrow \mathbb{T} = \mathbb{R}/\mathbb{Z}$. Je noterai f au lieu de T les homéomorphismes, afin de réserver la majuscule aux relèvements :

Rappelez-vous, du chapitre 3, la définition du degré : si f est un homéo de $\mathbb{T}$, on appelle relèvement de f à $\mathbb{R}$, une application $F : \mathbb{R} \longrightarrow \mathbb{R}$ telle que $\forall x \in \mathbb{R}, \pi(F(x)) = f(\pi(x))$. Alors, pour tout $x \in \mathbb{R}$, il existe $n(x) \in \mathbb{Z}$ tel que $F(x+1) = F(x) + n(x)$. Par continuité de F, l'application $x \mapsto n(x)$ est localement constante, et par connexité de $\mathbb{R}$ elle est constante.

Exercice 81. *Montrer, en utilisant l'injectivité de F, que $n = n(x) = \pm 1$.*

On dit que f préserve (resp. renverse) l'orientation si $n = 1$ (resp. $n = -1$). Cela revient à dire que F est croissante (resp. décroissante).

Dans la suite on considèrera des homéos préservant l'orientation, tous les résultats s'appliquent à des homéos renversant l'orientation, en composant par $x \mapsto -x$.

Exercice 82. *Soit x_n une suite réelle telle que $\forall n, m, x_{n+m} \leq x_n + x_m$. Montrer que x_n/n converge. Montrer que la conclusion est encore valable sous les hypothèses plus faible suivantes : $\forall n, m, x_{n+m} \leq x_n + x_m + Cste$, ou $\forall n, m, x_{n+m} \leq x_{n+Cste} + x_{m+Cste}$. Généraliser au cas d'une fonction $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ sous-additive ($f(x+y) \leq f(x) + f(y)$).*

Solution. Soit $l = \liminf x_n/n$ (l peut être $-\infty$). Soit $\phi : \mathbb{N} \to \mathbb{N}$ strictement croissante, telle que $x_{\phi(n)}/\phi(n) \to l$. Soit $\epsilon > 0$, et soit N tel que $\forall n \geq N$, on a $x_{\phi(n)}/\phi(n) \leq l + \epsilon$. Soit $N_1 \geq \phi(N)$ tel que pour tout $n \leq \phi(N)$, on a $x_n/N_1 \leq \epsilon$. Alors pour tout $n \geq N_1$, on a, avec la division euclidienne $n = q\phi(N) + r$, $x_n \leq q x_{\phi(N)} + x_r$, de sorte que

$$\frac{x_n}{n} \leq \frac{x_{\phi(N)}}{\phi(N)} + \frac{x_r}{n} \leq l + 2\epsilon.$$

Pour les variantes, remplacer x_n par $x_n + C$, puis par x_{n+2C}.

Proposition 7.1.1. *Soit f un homéo de $\mathbb{T}$ préservant l'orientation et F un relèvement de f à $\mathbb{R}$. Alors la limite*

$$\tau(f) := \lim_{n\to\infty} \frac{1}{n}(F^n(x) - x)$$

1. *existe pour tout $x \in \mathbb{R}$*
2. *ne dépend pas de x*
3. *ne dépend du choix de F qu'à un entier près : si $F1$, F_2 sont des relèvements de f, alors $\tau(F_1) - \tau(F_2) \in \mathbb{Z}$*
4. *est rationnelle, de la forme p/q, si f admet un point périodique de période q.*

Cette proposition justifie qu'on donne un nom à τ :

Définition 7.1.2. *On appelle nombre de rotation de f l'élément suivant de $\mathbb{T}$: $\tau(f) := \pi(\tau(F))$ (qui est bien défini grâce au point 3. ci-dessus).*

Proof. (de la proposition 7.1.1 ; cf. prop 11.1.1 du Katok).

Notons tout d'abord que F est strictement croissante, et $F(x+1) = F(x) + 1$ pour tout $x \in \mathbb{R}$, donc si $k \leq x < k+1$ avec $k \in \mathbb{Z}$, on a

$$\forall n \in \mathbb{N}, F^n(k) = F^n(0) + k \leq F^n(x) < F^n(0) + k + 1 \tag{7.1}$$

donc, pour tout $x \in \mathbb{R}$, on a

$$F^n(0) - 1 \leq F^n(x) - x < F^n(0) + 1$$

d'où, pour tous $x, y \in \mathbb{R}$,

$$|(F^n(x) - x) - (F^n(y) - y)| \leq 2.$$

Considérons alors la suite $u_n(x) = F^n(x) - x$, on a

$$u_{n+m}(x) = u_n(F^m(x)) + u_m(x) \leq u_n(x) + u_m(x) + 2$$

ce qui, avec l'exercice 82, démontre le point 1.

Le point 2. découle alors du fait que pour tous $x, y \in \mathbb{R}$, pour tout $n \in \mathbb{N}$, on a $|u_n(x) - u_n(y)| \leq 2$.

Le point 3. vient du fait que deux relèvements diffèrent d'un entier. En effet, si F_1, F_2 sont deux relèvements de f, tels que $F_1(x) = F_2(x) + k$, avec $k \in \mathbb{Z}$, pour tout x, alors $F_1^n(x) = F_2^n(x) + nk$ pour tout n : par récurrence,

$$F_1^{n+1}(x) = F_1(F_2^n(x) + nk) = F_1(F_2^n(x)) + nk = F_2(F_2^n(x) + k + nk.$$

Pour le point 4., supposons que x est périodique de période q, alors on a $F^q(x) = x + p$, où p est un entier, puisque $F^q(x)$ se projette sur x. Donc, pour tout $n \in \mathbb{N}$, on a $F^{nq}(x) = x + np$, de sorte que la suite $\frac{1}{nq}(F^{nq}(x) - x)$ tend vers p/q. Ceci démontre le point 4. □

Exercice 83. *Calculer le nombre de rotation d'une rotation !*

Proposition 7.1.3. *Le nombre de rotation est invariant par conjugaison topologique.*

Proof. Soit g un homéomorphisme de $\mathbb{T}$, supposons-le de degré 1 (l'autre cas est identique), et choisissons un relèvement G de g à $\mathbb{R}$ tel que $0 \leq G(0) < 1$, et $-1 \leq G^{-1}(0) < 0$.

On veut montrer que $\tau(G^{-1} \circ F \circ G) = \tau(F)$. Il suffit de montrer que $|G^{-1} \circ F^n \circ G(x) - F^n(x)|$ est borné indépendamment de x.

On a, pour tout $k \leq x < k + 1$ $(x \in \mathbb{R}, k \in \mathbb{Z})$,

$$\begin{gathered} k \leq G(0) + k \leq G(x) \leq G(0) + k + 1 \leq k + 2 \\ k - 1 \leq G^{-1}(0) + k \leq G^{-1}(x) \leq G^{-1}(0) + k + 1 \leq k + 1 \end{gathered}$$

d'où, avec (7.1), en notant $k_n = \lfloor F^n(0) \rfloor$,

$$\begin{gathered} k_n + k \leq F^n(0) + k \leq F^n(G(x)) \leq F^n(0) + k + 2 \leq k_n + k + 3 \\ k_n + k - 1 \leq G^{-1} \circ F^n \circ G(x) \leq k_n + k + 3 \\ F^n(0) + k - 2 \leq G^{-1} \circ F^n \circ G(x) \leq F^n(0) + k + 3 \end{gathered}$$

dont il résulte que $|G^{-1} \circ F^n \circ G(x) - F^n(x)| \leq 4$ pour tout x, et la proposition. □

Proposition 7.1.4. *(réciproque du point 4. de la proposition 7.1.1) Si $\tau(f)$ est rationnel, de la forme p/q, alors f admet un point périodique de période q.*

Proof. Observons tout d'abord que pour tout $k \in \mathbb{N}$, on a $\tau(f^k) = k\tau(f)$ mod 1. En effet, si F est un relèvement de f,

$$\lim_{n\to\infty} \frac{F^{kn}(0)}{n} = k \lim_{n\to\infty} \frac{F^{kn}(0)}{kn} = k\tau(F).$$

Par conséquent, $\tau(f) = p/q$ équivaut à $\tau(f^q) = 0$, et comme "f admet un point périodique de période q" équivaut à "f^q admet un point fixe", on est ramené à prouver que $\tau(f) = 0$ implique l'existence d'un point fixe de f.

Par contraposée, supposons que f n'a pas de point fixe. Cela signifie que pour aucun $x \in \mathbb{R}$, on n'a $F(x) - x \in \mathbb{Z}$. La fonction $x \mapsto F(x) - x$, qui est 1-périodique, admet donc un maximum $M < 1$ et un minimum $m > 0$. Alors pour tout $x \in \mathbb{R}$, $n \in \mathbb{N}$,

$$nm \leq F^n(x) - x = \sum_{k=1}^{n} F^k(x) - F^{k-1}(x) \leq nM$$

ce qui entraîne que $m \leq \tau(f) \leq M$, donc $\tau(f) \neq 0$. □

Exercice 84. *Construire un exemple où f a un nombre de rotation rationnel, mais f n'est pas une rotation (et même où f n'a qu'un seul point périodique).*

Indication: penser à $f(x) = x^2$.

Exercice 85. *Si $\tau(f)$ est rationnel, alors tous les points périodiques de f ont la même période.*

Exercice 86. *L'application $f \mapsto \tau(f)$ est continue pour la topologie C^0 (la topologie de la convergence uniforme).*

Solution : Soit f un homéomorphisme du cercle, et prenons un encadrement de la forme $p/q < \tau(f) < (p+1)/q$ de $\tau(f)$ par des rationnels, de sorte que f n'a pas de point périodique de période q. Alors, en relevant f à $\mathbb{R}$ en F, $F^q(x) - x$ n'est jamais entier, donc il existe un entier (forcément égal à p, par définition du nombre de rotation) tel que pour tout $x \in \mathbb{R}$, $p < F^q(x) - x < p + 1$. La fonction $x \mapsto F^q(x) - x$ est 1-périodique, donc elle admet un maximum $< p + 1$, et un minimum $< p$. Prenons maintenant g, qu'on relève en G, assez proche de f, dans la topologie C^0, pour que $p < G^q(x) - x < p + 1$, on a alors $p/q < \tau(g) < (p + 1)/q$.

L'application $f \mapsto \tau(f)$ explique le mot squelette que j'ai employé dans l'introduction de ce chapitre : elle permet de relier un homéo quelconque à une rotation.

Remarque 7.1.5. *Il est intéressant de considérer, pas seulement $\tau(f)$, mais $\tau(R_\theta \circ f)$, R_θ étant la rotation d'angle θ (peut faire l'objet d'un **exposé**). C'est beaucoup plus intéressant que ça en a l'air, vous êtes probablement en train de penser "ben quoi, $\tau(R_\theta \circ f) = \tau(f) + \theta$", et effectivement c'est le cas si f est une rotation, mais en général, pas du tout !*

7.2 La classification de Poincaré

7.2.1 Nombre de rotation rationnel

Dans ce paragraphe nous supposons que $\tau(f) = p/q$, avec p et q entiers et premiers entre eux. On a vu à la section précédente qu'alors f admet un point périodique x_0 de période q.

Lemme 7.2.1. *Il n'y a que deux types d'orbite non-périodique possibles pour f :*

- *si f a exactement une orbite périodique, alors tout point du cercle est hétérocline à deux points de l'orbite périodique*
- *si f a plus d'une orbite périodique, alors tout point du cercle est hétérocline à deux points d'orbites périodiques distinctes.*

Proof. Soit x un point périodique, et soit F un relèvement de f à R. Alors $F^q - p$ envoie l'intervalle $[x, x+1]$ dans lui-même, et on est ramené à un exo classique sur les applications de l'intervalle. □

Question : est-il vrai que si toute orbite est périodique, alors f est topologiquement conjuguée à une rotation (qui ne peut être que $R_{p/q}$, d'après la Proposition 7.1.3) ? équivalence orbitale, oui clairement, mais conjugaison ?

Voir les exos de la section 11.2 pour les classes de conjugaison, quand il y a des points non-périodiques. **Exposé** possible ?

7.2.2 Nombre de rotation irrationnel

Dans ce paragraphe nous supposons que $\tau(f) = \tau$, avec τ irrationnel. Le lemme suivant dit essentiellement que les points de l'orbite de x_0 sont disposés, sur le cercle, dans le même ordre que les points d'une orbite de la rotation d'angle τ.

Lemme 7.2.2. *Soient F un relèvement de f à $\mathbb{R}$, $x \in \mathbb{R}$, et $n_1, n_2, m_1, m_2 \in \mathbb{Z}$, on a*

$$n_1\tau + m_1 < n_2\tau + m_2 \Longleftrightarrow F^{n_1}(x) + m_1 < F^{n_2}(x) + m_2.$$

Proof. Observons tout d'abord que, n_1, n_2, m_1, m_2 étant fixés, l'expression $p(x) = F^{n_1}(x) + m_1 - F^{n_2}(x) - m_2$ ne s'annule jamais : si c'était le cas, on en déduirait l'existence d'un point périodique, ce qui contredit l'irrationnalité du nombre de rotation. En particulier, elle ne change jamais de signe, donc si l'inégalité

$$F^{n_1}(x) + m_1 < F^{n_2}(x) + m_2 \tag{7.2}$$

est vérifiée pour un x, elle est vérifiée pour tout x.

A présent, démontrons le sens $\Longleftarrow$ du lemme. Soient $x \in \mathbb{R}$, et $n_1, n_2, m_1, m_2 \in \mathbb{Z}$, tels que (7.2) est vérifiée. Posons $y = F^{-n_2}(0)$, d'après l'observation initiale, (7.2) est vérifiée pour y, d'où

$$F^{n_1 - n_2}(0) < m_2 - m_1$$

puis, F étant croissante,

$$\forall k \in \mathbb{N},\ F^{k(n_1-n_2)}(0) < k(m_2 - m_1).$$

Par conséquent

$$\tau = \lim_{k\to\infty} \frac{F^{k(n_1-n_2)}(0)}{k(n_2 - n_1)} < \frac{m_1 - m_2}{n_2 - n_1}$$

d'où $n_1\tau + m_1 < n_2\tau + m_2$.

Pour l'autre sens, procédons par contraposée, et démontrons

$$n_1\tau + m_1 \geq n_2\tau + m_2 \Longleftarrow F^{n_1}(x) + m_1 \geq F^{n_2}(x) + m_2.$$

Par l'observation initiale et l'irrationnalité de τ, l'égalité est impossible, des deux côtés de la flèche. On raisonne alors comme dans l'autre sens en remplaçant $<$ par $>$. □

Proposition 7.2.3. *Pour tout point x du cercle, l'ensemble ω-limite $\omega(x)$ ne dépend pas de x, et $\omega(x)$ est soit le cercle tout entier, soit d'intérieur vide et sans point isolé.*

Remarque. Si vous vous demandez comment on peut être fermé, d'intérieur vide et sans point isolé, pensez à l'ensemble de Cantor qu'on a déjà rencontré en étudiant l'expansivité.

D'abord un

Lemme 7.2.4. *Soient $x \in \mathbb{T}$, $m, n \in \mathbb{Z}$, et I un des deux intervalles sur le cercle d'extrémités $f^n(x)$ et $f^m(x)$. Alors toute orbite positive rencontre I.*

Proof. Il suffit de montrer que

$$\bigcup_{k\in\mathbb{N}} f^{-k}(I) = T.$$

Soit , pour tout $k \in \mathbb{N}$,

$$I_k = f^{-k(n-m)}(I) = \left[f^{-(k-1)n+km}(x), f^{-kn+(k+1)m}(x)\right].$$

Alors

$$I_{k+1} = f^{-(k+1)(n-m)}(I) = \left[f^{-kn+(k+1)m}(x), f^{-(k+1)n+(k+2)m}(x)\right]$$

donc les intervalles I_k et I_{k+1} ont une extrémité commune, par conséquent leur réunion est encore un intervalle. Donc la réunion de tous les I_k, sur $k \in \mathbb{N}$, est soit le cercle tout entier, soit un intervalle, mais dans ce dernier cas cela signifie qu'une des suites $f^{-(k-1)n+km}(x)$ ou $f^{-kn+(k+1)m}(x)$ converge, quand $k \to \infty$. Mais alors (toujours l'exo classique du lycée "$f^n(x)$ converge vers y, alors y est point fixe de f") la limite serait point fixe de $f^{-(n-m)}$, donc point périodique de f, ce qui contredit l'irrationalité du nombre de rotation. □

Proof. (de la proposition 7.2.3) Démontrons d'abord que $\omega(x)$ ne dépend pas de x. Soient $x, y \in \mathbb{T}$, et $z \in \omega(x)$. Il existe donc une suite d'entiers k_n strictement croissante telle que $f^{k_n}(x) \to z$. D'après le lemme 7.2.4, pour tout $n \in \mathbb{N}$, il existe $l_n \in \mathbb{N}$ tel que

$$f^{l_n}(y) \in \left[f^{k_n}(x), f^{k_{n+1}}(x)\right]$$

(on prend le petit intervalle, donc sa longueur tend vers zéro). Mais alors, par les gendarmes, $f^{l_n}(y)$ tend vers z, ce qui montre que $z \in \omega(y)$, donc $\omega(x) \subset \omega(y)$. En échangeant les rôles de x et y, on obtient que $\omega(x) = \omega(y)$.

Notons maintenant Ω l'ensemble limite commun à tous les points du cercle, et observons qu'Ω est le seul fermé invariant minimal (au sens où il ne contient pas de fermé invariant non vide et distinct de lui-même) : en effet, pour tout x dans Ω, l'adhérence de l'orbite de x contient $\omega(x) = \Omega$.

Maintenant, observons que la frontière topologique $\partial\Omega$ de Ω (l'ensemble des points de Ω dont tout voisinage rencontre le complémentaire de Ω) est un fermé invariant de Ω, donc ou bien $\partial\Omega = \emptyset$, auquel cas $\Omega = \mathbb{T}$, ou bien $\partial\Omega = \Omega$, auquel cas Ω est d'intérieur vide.

Il reste à montrer que Ω n'a pas de point isolé. Pour tout $x \in \Omega$, on a $\omega(x) = \Omega$, donc x est valeur d'adhérence de la suite $f^n(x)$. Si x était isolé dans Ω, x serait alors égal à $f^n(x)$ pour un certain n, c'est à dire périodique, ce qui contredit l'irrationalité du nombre de rotation. □

Théorème 7.2.5. *Sous les hypothèses de ce paragraphe, si f est transitive, alors f est topologiquement conjuguée à la rotation R_τ. Si f n'est pas transitive, alors on a juste une semi-conjugaison : il existe une application continue h du cercle dans lui-même (en général pas bijective) telle que $R_\tau \circ h = h \circ f$.*

Proof. Soient F un relévement de f à $\mathbb{R}$, $x \in \mathbb{R}$, et $\tau := \tau(f)$. Notons $B = \{F^n(x) + m \ : \ n, m \in \mathbb{Z}\}$ le relèvement à $\mathbb{R}$ de l'orbite par f de $\pi(x)$. Définissons une application par

$$\begin{aligned} H : B \quad &\longrightarrow \quad \mathbb{R} \\ F^n(x) + m \quad &\longmapsto \quad n\tau + m. \end{aligned}$$

Notons que l'application H passe au quotient en une application h définie sur l'orbite de $\pi(x)$, à valeur dans le cercle, telle que $R_\tau \circ h = h \circ f$.

Par la classification des sous-groupes de $(\mathbb{R}, +)$ (ou, ce qui revient au même, la minimalité des rotations irrationnelles), $H(B)$ est dense dans $\mathbb{R}$, donc $h(\pi(B))$ est dense dans le cercle.

Montrons maintenant qu'on peut étendre continûment h à tout le cercle. On commence par l'étendre à l'adhérence de l'orbite de $\pi(x)$. Soit $y \in \overline{B}$. D'après le lemme 7.2.2, l'application H est monotone, donc admet des limites à droites et à gauche en y. Si ces limites ne coïncidaient pas, toujours par monotonie de H, H éviterait tout l'intervalle compris entre la limite à droite et la limite à gauche, ce qui contredit la densité de $H(B)$. On peut donc prolonger H par continuité à $\overline{B}$.

On prolonge ensuite H à $\mathbb{R}$ tout entier, en prenant H constante sur les composantes connexes du complémentaire de $\overline{B}$: si $[a, b]$ est une composante connexe du complémentaire, on a $H(a) = H(b)$, sans quoi H éviterait tout l'intervalle compris entre $H(a)$ et $H(b)$, donc le prolongement défini en prenant H constante sur $[a, b]$ est continu.

Le prolongement vérifie encore $H(F(x)) = H(x) + \tau$, et passe au quotient en une application h définie sur le cercle, à valeur dans le cercle, telle que $R_\tau \circ h = h \circ f$.

Cette application est la semi-conjugaison demandée ; pour achever la preuve du théorème il suffit de voir que cette semi-conjugaison est en fait une conjugaison quand l'orbite de x est dense. En effet dans ce cas $\overline{B} = \mathbb{R}$, donc h est une bijection continue, donc un homéomorphisme du cercle (exercice : une bijection continue d'un compact est un homéomorphisme). $\square$

Comme dans le cas rationnel, on peut classifier les types topologiques d'orbite, et on en déduit le

Corollaire 7.2.6. *Tout homéomorphisme f du cercle de nombre de rotation irrationnel est uniquement ergodique, et la mesure invariante admet une densité par rapport à la mesure de Lebesgue si et seulement si f est conjugué à une rotation.*

Proof. (esquisse) : toute mesure invariante est envoyée par la semi-conjugaison sur une mesure invariante de la rotation. $\square$

Question : les rotations irrationnelles sont-elles structurellement stables?

Réponse : non, c'est le théorème de Denjoy (exposé ?). Par contre on a une version plus faible de la stabilité structurelle, où on remplace C^1 par C^2.

Exercice 87. *Classifier la dynamique du pendule en restriction à chaque niveau d'énergie. Montrer que tous les nombres de rotations possibles sont réalisés par au moins un niveau d'énergie.*

Chapter 8

Entropie topologique

8.1 Définition

But de cette section : quantifier (au moyen d'un seul nombre) la façon dont les orbites divergent. On commence par modifier la distance de façon à tenir compte de la divergence des orbites: :

$$d_n^T(x,y) = \max_{i=0,\dots,n-1} d(T^i(x), T^i(y))$$

Exercice 88. *Montrer que cette formule définit bien une distance, et que cette distance induit sur X la même topologie que d.*

Cet exercice entraîne que, (X,d) étant compact, (X, d_n^T) l'est aussi, pour tout n. Donc pour tout $\epsilon > 0$, on peut recouvrir X par un nombre fini de d_n^T-boules de rayon $\leq \epsilon$.

On définit $S(T,\epsilon,n)$ comme étant le nombre minimum de d_n^T-boules de rayon $\leq \epsilon$ nécessaire pour recouvrir X. Ensuite on regarde le taux de croissance exponentiel de $S(T,\epsilon,n)$:

$$h(T,\epsilon) := \limsup_{n\to+\infty} \frac{1}{n} \log S(T,\epsilon,n)$$

Le choix du taux de croissance exponentiel est un peu arbitraire, il se justifie par l'habitude de considérer que croissance rapide = croissance exponentielle.Il sera également justifié, a posteriori, par les résultats.

Définition 8.1.1 (Adler, Konheim, McAndrew, 1969)**.** *On appelle entropie topologique du système (X,T) la limite $h(T) = \lim_{\epsilon\to 0} h(T,\epsilon)$.*

Ainsi dit, il n'est même pas évident que ça existe !

Exercice 89. *Montrer l'existence de la limite quand ϵ tend vers zéro.*

Indication : montrer qu'on a une fonction croissante de ϵ.

Vous observerez que je n'ai pas indiqué de dépendance en la distance dans la notation. La raison en est :

Exercice 90. *Montrer que si une autre distance d' induit sur X la même topologie que d, alors les entropies topologiques de T, calculées au moyen de d ou d', sont les mêmes.*

Cet exercice justifie le nom d'entropie topologique.

Il y a une autre façon de définir l'entropie topologique, qui nous sera utile par la suite. On définit $N(T, \epsilon, n)$ comme le cardinal maximal d'une partie de X d_n^T, ϵ-séparée (δ-séparée signifie que deux points distincts sont distants d'au moins δ pour la distance d_n^T. L'existence, et la finitude, de $N(T, \epsilon, n)$ résulte de la compacité de X).

Exercice 91. *Montrer que*

$$h(T) = \lim_{\epsilon \to 0} \limsup_{n \to +\infty} \frac{1}{n} \log N(T, \epsilon, n).$$

Indication : montrer que $S(T, \epsilon, n) \leq N(T, \epsilon, n) \leq S(T, \frac{\epsilon}{2}, n)$

Question : est-il vrai qu'une application de l'intervalle, d'entropie positive, possède des orbites périodiques de toutes périodes ?

8.2 Propriétés

Proposition 8.2.1. *1. L'entropie d'un sous-système est toujours plus petite que l'entropie du système total : si $Y \subset X$ est un fermé invariant par T, alors $h(T_{|Y}) \leq h(T)$.*

2. Entropie d'une réunion de sous-systèmes : si $X = \cup_{i=1}^n X_i$, où les X_i sont des fermés invariants par T, alors

$$h(T) = \max_{i=1,\ldots,n} h(T_{|X_i})$$

3. l'entropie est comme un logarithme, elle change les produits en somme: pour tout $n \in \mathbb{N}$, $h(T^n) = nh(T)$.

4. Invariance par conjugaison : si f est un homéomorphisme de X dans Y, alors $h(X, T) = h(Y, f \circ T \circ f^{-1})$.

Proof. 1-tout recouvrement ouvert de X est un recouvrement de Y, donc $S(T, \epsilon, n) \geq S(T_{|Y}), \epsilon, n)$ pour tous ϵ, n.

2-d'après 1-, on a

$$h(T) \geq \max_{i=1,\ldots,n} h(T_{|X_i})$$

D'autre part, une réunion de recouvrements de $X_1, \ldots, X_n$, est un recouvrement de X, on a donc

$$S(T, \epsilon, k) \leq \sum_{i=1}^{n} S(T_{|X_i}), \epsilon, k) \forall \epsilon, k.$$

Par conséquent, pour tout k, ϵ, il existe $i \in \{1, \ldots, n\}$ tel que

$$\frac{1}{n} S(T, \epsilon, k) \leq S(T_{|X_i}), \epsilon, k) \tag{8.1}$$

et comme il y a une infinité de k, ϵ et seulement n indices i possibles, il existe un i tel que (8.1) vaut pour une infinité de valeurs de k, et pour une suite tendant vers 0 de valeurs de ϵ. On a alors

$$\log S(T, \epsilon, k) \leq \log S(T_{|X_i}), \epsilon, k) - \log n$$

et on conclut avec la gymnastique habituelle (diviser par k, prendre la limsup quand k tend vers l'infini, prendre la limite quand ϵ tend vers zéro).

3-remarquons tout d'abord que pour tous x, y, n, on a

$$d_k^{T^n}(x, y) \leq d_{nk}^T(x, y)$$

puisqu'à gauche on prend un un max sur $i = 0, n, \ldots, kn$ tandis qu'à droite on prend le max sur $0, 1, 2, \ldots, nk$ (si on prend le max sur un ensemble plus grand, le max est plus grand). Donc

$$\frac{1}{k} \log S(T^n, \epsilon, k) \leq n \frac{1}{nk} \log S(T, \epsilon, nk)$$

donc $h(T^n, \epsilon) \leq nh(T, \epsilon)$, puis $h(T^n) \leq nh(T)$ en faisant tendre ϵ vers zéro.

Réciproquement, puisque d_{nk}^T et $d_k^{T^n}$ définissent la même topologie, l'identité

$$\left(X, d_k^{T^n}\right) \longrightarrow \left(X, d_{nk}^T\right)$$

est continue, et comme X est compact, elle est uniformément continue. Donc pour tout $\epsilon > 0$, il existe $\delta(\epsilon) > 0$, tel que toute boule de rayon $\delta(\epsilon)$ pour $d_k^{T^n}$ est contenue dans une boule de rayon ϵ pour d_{nk}^T.

Donc tout recouvrement de X par des boules de rayon $\delta(\epsilon)$ pour $d_k^{T^n}$ induit un recouvrement de X par des boules de rayon ϵ pour d_{nk}^T, donc $S(T, \epsilon, nk) \leq S(T^n, \delta(\epsilon), k)$. Par conséquent $h(T^n, \epsilon) \geq nh(T, \epsilon)$, puis $h(T^n) \geq nh(T)$.

4-Exercice, en utilisant l'exercice 90. □

L'invariance par conjugaison s'utilise en général négativement, pour montrer que deux systèmes ne sont pas topologiquement conjugués.

Exercice 92 (théorème de Kolmogorov). *Montrer que le décalage sur les suites de* 0 *et de* 1 *n'est pas conjugué topologiquement (i.e. n'est pas le même système dynamique, même en changeant de lunettes) au décalage sur les suites de* 0, 1, *et* 2.

Exercice 93. *Calculer l'entropie d'une rotation.*

En fait on peut dire mieux :

Exercice 94. *Tout homéomorphisme du cercle est d'entropie topologique nulle.*

Indication : utiliser l'exercice précédent, le chapitre précédent, et le théorème 4.

Exercice 95. *Calculer l'entropie du décalage (et des applications linéaires dilatantes sur le cercle ?)*

Exercice 96. *Calculer l'entropie du sous-décalage sur les suites de* 0 *et de* 1 *où* 0 *est toujours suivi de* 1*.*

8.2.1 Entropie topologique et entropie en physique

On a dû vous dire en physique quand vous étiez petits, que si on a deux récipients contenant chacun un gaz, l'entropie du système "réunion des deux récipients" est la somme des entropies de chacun des récipients. La proposition suivante prouve que notre notion d'entropie n'est pas la même que celle des physiciens.

Proposition 8.2.2. *Si on a deux systèmes dynamiques* (X,T) *et* (Y,R) *(*X *et* Y *étant supposés disjoints comme ensembles), alors le système réunion :*

$$\begin{array}{rcl} T\cup R : X\cup Y & \longrightarrow & X\cup Y \\ x & \longmapsto & T(x) \text{ si } x\in X \\ & & R(x) \text{ si } x\in Y \end{array}$$

a pour entropie $\max(h(T),h(R))$*.*

Proof. Prenons K plus grand que le diamètre de X, et que le diamètre de Y. On munit $X\cup Y$ de la distance

$$\begin{array}{rcl} d(x,y) & = & d_X(x,y) \text{ si } x,y\in X \\ & & d_Y(x,y) \text{ si } x,y\in Y \\ & & K \text{ sinon .} \end{array}$$

Je vous laisse vérifier que cette distance définit bien la topologie de la réunion disjointe, en particulier X et Y sont ouverts et fermés. On a alors

$$\begin{array}{rcl} d_n^{T\cup R}(x,y) & = & d_{X,n}^T(x,y) \text{ si } x,y\in X \\ & & d_{Y,n}^R(x,y) \text{ si } x,y\in Y \\ & & K \text{ sinon .} \end{array}$$

On conclut alors avec la propriété 8.2.1, item 2. □

Au chapitre suivant nous parlerons de l'entropie des physiciens, qui est aussi celle de la théorie de l'information.

Chapter 9

Entropie mesurée

9.1 Entropie mesurée d'une partition

Reprenons l'exemple du décalage sur les suites de 0 et de 1. On interprète une telle suite comme un message qu'on voudrait déchiffrer, et le décalage comme l'opération de lecture, une lettre à la fois. On se demande si, à partir de ce qu'on a déjà lu, on est capable de prédire la prochaine lettre.

Dans le cas du "full shift" (le décalage sur l'ensemble de toutes les suites), si les lettres 0 et 1 sont équiprobables, on ne peut rien dire. Par contre, dans l'exemple où 0 est toujours suivi de 1, si on lit 0, on connaît la prochaine lettre. L'incertitude est donc plus grande dans le cas du "full shift".

Le but de cette section est de quantifier la notion d'incertitude. Pour cela, on se place dans le même cadre (X, d, T) que précédemment, mais on se donne en plus une mesure borélienne de probabilité μ sur X.

Définition 9.1.1. *On appelle partition mesurée de* (X, μ) *une collection finie ou dénombrable de parties mesurables de* X*,* $\mathcal{A} = \{A_i, i \in I\}$*, telle que* $\forall i, \mu(A_i) > 0$*,* $\mu(X \setminus \cup_i A_i) = 0$*, et* $\mu(A_i \cap A_j) = 0$ *pour* $i \neq j$*.*

Une partition mesurée est donc une vraie partition de X, à un ensemble de mesure nulle près.

Définition 9.1.2. *On appelle entropie de la partition mesurée* $\mathcal{A}$*, la quantité* $H(\mathcal{A}) = -\sum_i \mu(A_i) \log \mu(A_i)$*.*

convention : $0 \log 0 = 0$

interprétation : la quantité $-\log \mu(A_i)$ peut s'interpréter comme l'information gagnée en apprenant qu'on se trouve dans A_i : plus A_i est petit, et plus le fait d'être dans A_i donne une information précise sur notre position, ce qui correspond au fait que $-\log \mu(A_i)$ est grand. Inversement, si $\mu(A_i) = 1$, savoir qu'on est dans A_i ne nous apprend pas grand chose puisque μ-presque tout point est dans A_i, ce qui correspond au fait que $-\log \mu(A_i) = 0$. Jusque là vous vous demandez peut-être pourquoi on prend le log et pas n'importe quelle fonction f négative entre 0 et 1 telle que $f(1) = 0$, $\lim f(x) = -\infty$ et $\lim xf(x) = 0$, mais vous verrez dans la suite que le choix du log n'est pas gratuit.

L'entropie est l'information gagnée, en moyenne (espérance, dans le language des variables aléatoires), en apprenant dans quel A_i on se trouve. Si on voit la partition comme n rangement d'un tas de papiers, l'entropie est l'efficacité du rangement. A votre avis, si vous avez un gros tas de papiers, et N tiroirs, quelle est la façon de répartir les papiers dans les tiroirs qui minimisera votre temps moyen de recherche pour trouver un papier ? Si vous avez répondu "il faut en mettre autant dans chaque tiroir", votre intution est correcte, comme le prouve l'exercice suivant.

Exercice 97. *montrer que parmi toutes les partitions finies à N éléments, celles dont les éléments sont équiprobables maximisent l'entropie (qui vaut alors $\log N$).*

indication : concavité du log

Proposition 9.1.3. *concavité de l'entropie par rapport à la mesure : si λ est une autre mesure, $\forall t \in [0,1]$, $tH_\mu(\mathcal{A}) + (1-t)H_\lambda(\mathcal{A}) \leq H_{t\mu+(1-t)\lambda}(\mathcal{A})$.*

Proof. La fonction $\phi(x) := -x \log x$ est concave sur $[0,1]$, donc pour tout i,

$$t\phi(\mu(A_i)) + (1-t)\phi(\lambda(A_i)) \leq \phi(t\mu(A_i) + (1-t)\lambda(A_i))$$

puis on somme sur i. $\square$

9.2 Entropie conditionnelle

Rappel pour les non-probabilistes parmi vous : la probabilité de A sachant B, notée $\mu(A|B)$, est $\mu(A \cap B)/\mu(B)$.

Définition 9.2.1. *$\mathcal{A} \leq \mathcal{B}$ ($\mathcal{B}$ raffine $\mathcal{A}$) : $\forall j \in J, \exists i \in I, B_j \subset A_i$*

Définition 9.2.2. *conjonction de $\mathcal{A}$ et $\mathcal{B}$: $\mathcal{A} \vee \mathcal{B} = \{A_i \cap B_j : (i,j) \in I \times J\}$*

Interprétation : plus petit raffinement commun de $\mathcal{A}$ et $\mathcal{B}$

Définition 9.2.3. *entropie de $\mathcal{A}$ conditionnellement à $\mathcal{B}$:*

$$\begin{aligned} H(\mathcal{A}|\mathcal{B}) &= -\sum_{i,j} \mu(A_i \cap B_j) \log \frac{\mu(A_i \cap B_j)}{\mu(B_j)} \\ &= -\sum_j \mu(B_j) \sum_i \mu(A_i|B_j) \log \mu(A_i|B_j) \end{aligned}$$

C'est l'information gagnée (=réduction d'incertitude), en moyenne, en apprenant dans quel A_i on se trouve, lorsqu'on sait déjà dans quel B_j on se trouve.

9.2.1 Propriétés de l'entropie conditionnelle

1. $0 \leq H(\mathcal{A}|\mathcal{B}) \leq H(\mathcal{A})$
2. $H(\mathcal{A}|\mathcal{B}) = H(\mathcal{A})$ si et seulement si $\mathcal{A}$ et $\mathcal{B}$ sont indépendantes
3. $H(\mathcal{A}|\mathcal{B}) = 0$ si et seulement si $\mathcal{A} \leq \mathcal{B}$
4. $\mathcal{C} \geq \mathcal{B}$ entraîne $H(\mathcal{A}|\mathcal{C}) \leq H(\mathcal{A}|\mathcal{B})$
5. $H(\mathcal{A} \vee \mathcal{B}) = H(\mathcal{A}) + H(\mathcal{B}|\mathcal{A})$
6. $H(\mathcal{A} \vee \mathcal{B}) \leq H(\mathcal{A}) + H(\mathcal{B})$

Interprétation et commentaires

Si on veut juste définir l'entropie mesurée, on n'a besoin que de 1, 5 et 6.

1-pour la première inégalité : pas d'interprétation (l'incertitude ne peut pas être négative : on ne peut pas être plus que certain ?). Pour la deuxième inégalité : savoir dans quel élément de $\mathcal{B}$ on se trouve ne peut que réduire, et en tous cas pas augmenter, l'incertitude.

2-Dire que $\mathcal{A}$ et $\mathcal{B}$ sont indépendantes, revient exactement à dire que savoir dans quel élément de $\mathcal{B}$ on se trouve n'apporte aucune information relative à $\mathcal{A}$.

3-si, une fois qu'on sait dans quel élément de $\mathcal{B}$ on est, il n'y a plus aucune incertitude par rapport à $\mathcal{A}$, cela signifie que que les éléments de $\mathcal{B}$ sont contenus dans des éléments de $\mathcal{A}$.

4-si $\mathcal{C}$ raffine $\mathcal{B}$, savoir dans quel élément de $\mathcal{C}$ on se trouve apporte plus d'information que savoir savoir dans quel élément de $\mathcal{B}$ on se trouve (ou bien : une information plus précise réduit davantage l'incertitude).

5 et 6-en apprenant dans quel élément de $\mathcal{A} \vee \mathcal{B}$ on se trouve, on apprend à la fois dans quel élément de $\mathcal{A}$, et dans quel élément de $\mathcal{B}$, on se trouve. C'est là qu'avoir pris le log dans la définition de l'entropie est fondamental.

Preuves

1. pour la première inégalité, on observe que la fonction $\phi(x) = -x \log x$ est ≥ 0 sur $[0, 1]$. Pour l'autre inégalité :

$$\begin{aligned} H(\mathcal{A}|\mathcal{B}) &= \sum_j \mu(B_j) \sum_i \phi(\mu(A_i|B_j)) \\ &= \sum_i \sum_j \mu(B_j) \phi(\mu(A_i|B_j)) \end{aligned}$$

or la fonction ϕ est concave, et $\sum_j \mu(B_j) = 1$, donc pour tout i,

$$\begin{aligned}
\sum_j \mu(B_j)\phi(\mu(A_i|B_j)) &\leq \phi\left(\sum_j \mu(B_j)\mu(A_i|B_j)\right) \qquad (9.1)\\
&= \phi\left(\sum_j \mu(A_i \cap B_j)\right)\\
&= \phi\left(\cup_j A_i \cap B_j\right) = \phi(A_i)
\end{aligned}$$

où la première égalité est la définition de la probabilité conditionnelle, la deuxième égalité vient du fait que la réunion est disjointe, et la troisième égalité vient du fait que les B_j partitionnent A_i. On a donc

$$H(\mathcal{A}|\mathcal{B}) \leq \sum_i \phi(A_i) = H(\mathcal{A}).$$

2. Si $\mathcal{A}$ et $\mathcal{B}$ sont indépendantes, $\mu(A_i \cap B_j) = \mu(A_i)\mu(B_j)$ pour tous i, j, donc

$$\begin{aligned}
H(\mathcal{A}|\mathcal{B}) &= -\sum_{i,j} \mu(A_i)\mu(B_j) \log \mu(A_i)\\
&= -\sum_i \mu(A_i) \log \mu(A_i) \left(\sum_j \mu(B_j)\right) = H(\mathcal{A})
\end{aligned}$$

en remarquant que le terme entre parenthèses vaut 1.

Réciproquement, si $H(\mathcal{A}|\mathcal{B}) = H(\mathcal{A})$, cela signifie que toutes les inégalités du 1- sont des égalités, en particulier (9.1), mais la fonction ϕ est strictement concave, donc l'égalité dans (9.1) entraîne que $\mu(A_i|B_j)$ ne dépend pas de j, soit m sa valeur, on a alors $\mu(A_i \cap B_j) = m\mu(B_j)$, et en sommant sur tous les B_j (qui partitionnent), on obtient $\mu(A_i|B_j) = \mu(A_i)$.

3. On a $H(\mathcal{A}|\mathcal{B}) = 0$ si et seulement si pour tous i, j, $\phi(\mu(A_i|B_j)) = 0$, c'est à dire $\mu(A_i|B_j) = 0$ ou 1, soit encore $B_j \subset A_i$, ou $A_i \cap B_j$ (à mesure nulle près bien sûr). Comme $\mu(B_j) > 0$ pour tout j, cela entraîne que pour tout j, il existe i, tel que $B_j \subset A_i$, c'est à dire $\mathcal{B}$ raffine $\mathcal{A}$.

4. On veut démontrer

$$\sum_k \mu(C_k) \sum_i \phi(\mu(A_i|C_k)) \leq \sum_j \mu(B_j) \sum_i \phi(\mu(A_i|B_j)). \qquad (9.2)$$

Observons que $\sum_i \phi(\mu(A_i|B_j))$ est l'entropie de la partition de B_j en les $A_i \cap B_j$, muni de la mesure conditionnelle $\mu_j(.) = \mu(. \cap B_j)/\mu(B_j)$.

On peut alors lui appliquer la deuxième inégalité du point 1., et on obtient

$$\begin{aligned}\sum_k \mu_j(C_k \cap B_j) \sum_i \phi\left(\mu_j(A_i|C_k)\right) &\leq \sum_i \phi\left(\mu_j(A_i)\right) \\ &= \sum_i \phi(\mu(A_i|B_j)) \qquad (9.3)\end{aligned}$$

mais comme $\mathcal{C} \geq \mathcal{B}$, on a, pour tout k, $\mu_j(C_k) = 0$ (si C_k n'est pas inclus dans B_j), ou $\mu_j(C_k) = \mu(C_k)/\mu(B_j)$ (si $C_k \subset B_j$), et

$$\begin{aligned}\mu_j(A_i|C_k) &= \frac{\mu_j(A_i \cap C_k)}{\mu_j(C_j)} = \frac{\mu(A_i \cap B_j \cap C_k)}{\mu(B_j \cap C_k)} \\ &= \frac{\mu(A_i \cap C_k)}{\mu(C_k)} \text{ si } C_k \subset B_j \\ &= 0 \text{ sinon}\end{aligned}$$

et (9.3) devient

$$\sum_{k, C_k \subset B_j} \frac{\mu(C_k)}{\mu(B_j)} \sum_i \phi\left(\frac{\mu(A_i \cap C_k)}{\mu(C_k)}\right) \leq \sum_i \phi(\mu(A_i|B_j)).$$

Observons que

$$\frac{\mu(A_i \cap C_k)}{\mu(C_k)} = \mu(A_i|C_k),$$

et en sommant sur j, et multipliant par $\mu(B_j)$, on obtient (9.2).

5.

$$\begin{aligned}H(\mathcal{B}|\mathcal{A}) &= -\sum_{i,j} \mu(A_i \cap B_j) \log \frac{\mu(A_i \cap B_j)}{\mu(A_i)} \\ &= -\sum_{i,j} \mu(A_i \cap B_j) \log \mu(A_i \cap B_j) + \sum_{i,j} \mu(A_i \cap B_j) \log \mu(A_i) \\ &= H(\mathcal{A} \vee \mathcal{B}) + \sum_i \log \mu(A_i) \left(\sum_j \mu(A_i \cap B_j)\right) \\ &= H(\mathcal{A} \vee \mathcal{B}) + \sum_i (\log \mu(A_i))\, \mu(A_i) \\ &= H(\mathcal{A} \vee \mathcal{B}) - H(\mathcal{A})\end{aligned}$$

6. c'est juste 5+1.

9.3 Entropie mesurée d'un système dynamique

On suppose désormais que T préserve la mesure.

$$\mathcal{A}_{T,n} = \mathcal{A} \vee T^{-1}\mathcal{A} \ldots \vee T^{-n+1}\mathcal{A}$$

Définition 9.3.1. *On appelle entropie de T relativement à la mesure μ et à la partition $\mathcal{A}$, et on note $h_\mu(T, \mathcal{A})$, la limite suivante :*

$$\lim_{n\to+\infty} \frac{1}{n} H(\mathcal{A}_{T,n})$$

Preuve de l'existence de la limite : on a

$$\mathcal{A}_{T,n+m} = \mathcal{A}_{T,n} \vee T^{-n}\left(\mathcal{A}_{T,m}\right)$$

donc d'après le point 6- ci-dessus,

$$\begin{aligned} H\left(\mathcal{A}_{T,n+m}\right) &\leq H\left(\mathcal{A}_{T,n}\right) + H\left(\mathcal{A}_{T,m}\right) \\ &= H\left(\mathcal{A}_{T,n}\right) + H\left(\mathcal{A}_{T,m}\right) \end{aligned}$$

donc la suite $H\left(\mathcal{A}_{T,n+m}\right)$ est sous-additive, et on conclut avec l'exo 82 :

Définition 9.3.2. *On appelle entropie de T relativement à la mesure μ, la quantité*

$$h_\mu(T) = \sup_{\mathcal{A}} h_\mu(T, \mathcal{A}).$$

C'est cette notion d'entropie qui est utilisée en physique ; ils ont une façon mystérieuse (pour moi) d'associer à un système physique (en équilibre thermodynamique ?) une mesure de probabilité (mesure de Gibbs ?).

Propriétés de l'entropie mesurée

1. $h_\mu(T, \mathcal{A}) \leq h_\mu(T, \mathcal{B}) + H(\mathcal{A}|\mathcal{B})$
2. $h_\mu(T^k) = kh_\mu(T)$ pour $k \in \mathbb{N}$
3. si λ est une autre mesure, $\forall t \in [0,1]$, $th_\mu(T) + (1-t)h_\lambda(T) \leq h_{t\mu+(1-t)\lambda}(T)$.

9.4 Principe variationnel

Voici le lien entre l'entropie mesurée et l'entropie topologique :

Théorème 9.4.1 (divers auteurs, fin 60's)**.** *Pour un homéomorphisme T d'un espace métrique compact X, on a $h(T) = \sup_\mu h_\mu(T)$, où le supremum est pris sur l'ensemble des mesures boréliennes de probabilité invariantes par T.*

Peut faire l'objet d'un exposé

Pour démontrer le théorème, nous avons besoin d'une machine qui fabrique des mesures de grande entropie :

Proposition 9.4.2. *Soient*

- (X, d) *un espace métrique compact,*
- T *un homéomorphisme de* X*,*
- $\epsilon > 0$*,*
- $E_n \subset X$ *une partie* (n, ϵ)*-séparée*
- ν_n *la mesure uniforme sur* E_n*, définie par*

$$\nu_n = \frac{1}{|E_n|} \sum_{x \in E_n} \delta_x$$

- *et* μ_n *la mesure sur* X *définie par*

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} T_*^i \nu_n$$

Alors il existe un point d'accumulation μ *de la suite* μ_n *pour la topologie faible*, tel que* μ *est* T*-invariant, et*

$$\limsup \frac{1}{n} \log |E_n| \leq h_\mu(T).$$

Proof. Soit $\mathcal{A}$ une partition mesurée finie de X, dont tous les éléments sont de diamètre $< \epsilon$. Montrons d'abord

$$\log |E_n| = H_{\nu_n}(\mathcal{A}_{T,n}). \tag{9.4}$$

Lemme 9.4.3. *Les éléments de* $\mathcal{A}_{T,n}$ *sont de diamètre* $< \epsilon$ *pour la distance* d_n^T*.*

Preuve du lemme : Supposons

$$\exists x_1, x_2 \in A_1 \cap T^{-1}(A_2) \cap \ldots \cap T^{-n+1}(A_n) \text{ tels que } d_n^T(x_1, x_2) \geq \epsilon,$$

alors il existe $i = 0, \ldots, n-1$ tel que $d(T^i(x_1), d(T^i(x_2) \geq \epsilon$, mais $x_1, x_2 \in T^{-i}(A_{i+1})$, donc $T^i(x_1), T^i(x_2) \in A_{i+1}$, qui est de diamètre $< \epsilon$. Cette contradiction achève la preuve du lemme.

Par conséquent, et puisque E_n est (n, ϵ)-séparée, un élément de $\mathcal{A}_{T,n}$ contient au plus un élément de E_n. Donc, pour $A \in \mathcal{A}_{T,n}$, on a $\nu_n(A) = 0$ ou $1/|E_n|$ selon que A contient, ou non, un élément de E_n. Donc

$$\begin{aligned} H_{\nu_n}(\mathcal{A}_{T,n}) &= \sum_{A \in \mathcal{A}_{T,n}} -\nu_n(A) \log \nu_n(A) \\ &= \sum_{A \in \mathcal{A}_{T,n}, A \cap E_n \neq \emptyset} \frac{1}{|E_n|} \log |E_n| \end{aligned}$$

et comme il y a exactement $|E_n|$ éléments de $\mathcal{A}_{T,n}$ qui contiennent un élément de E_n, on obtient (9.4).

Maintenant, prenons un entier q, $0 < q < n$. Nous allons montrer que

$$q \log |E_n| \leq H_{\mu_n}(\mathcal{A}_{T,q}) + 2q^2 \log |\mathcal{A}|. \tag{9.5}$$

Pour $0 \leq k < q$, définissons $a(k) := \lfloor (n-k)/q \rfloor$. On a donc

$$\{0, 1, \ldots, n-1\} = \{k + rq + i : 0 \leq r < a(k), 0 < i < q\} \cup S$$

où S est défini par

$$S = \{0, 1, \ldots, k, k + a(k)q + 1, \ldots, n-1\}.$$

Notons que $|S| \leq 2q$ puisque $k < q$ et $k + a(q)q \geq n - q$ (puisque $a(k) \geq (n-k)/q - 1$). On a, par définition de S,

$$\mathcal{A}_{T,n} = \left(\bigvee_{r=0}^{a(k)-1} T^{-rq-k}(\mathcal{A}_{T,q}) \right) \vee \left(\bigvee_{i \in S} T^{-i}(\mathcal{A}) \right).$$

Rappelons que d'après 9.2.1, item 6., on a $H(\mathcal{A} \vee \mathcal{B}) \leq H(\mathcal{A}) + H(\mathcal{B})$, donc

$$\log |E_n| = H_{\nu_n}(\mathcal{A}_{T,n}) \leq \sum_{r=0}^{a(k)-1} H_{\nu_n}\left(T^{-rq-k}(\mathcal{A}_{T,q})\right) + \sum_{i \in S} H_{\nu_n}(T^{-i}(\mathcal{A}))$$

et puisque pour tous $\mu, \mathcal{A}$ on a $H_\mu(\mathcal{A}) \leq \log |\mathcal{A}|$, et que $|S| \leq 2q$, on peut majorer le deuxième terme à droite par $2q \log |\mathcal{A}|$.

D'autre part

$$\forall r, k \ H_{\nu_n}\left(T^{-rq-k}(\mathcal{A}_{T,q})\right) = H_{T_*^{rq+k}\nu_n}(\mathcal{A}_{T,q})$$

d'où

$$\log |E_n| \leq \sum_{r=0}^{a(k)-1} H_{T_*^{rq+k}\nu_n}(\mathcal{A}_{T,q}) \ + \ 2q \log |\mathcal{A}|$$

puis, en sommant sur $k = 0, \dots, q-1$,

$$q \log |E_n| \leq \sum_{k=0}^{q-1} \sum_{r=0}^{a(k)-1} H_{T_*^{rq+k}\nu_n}(\mathcal{A}_{T,q}) \;+\; 2q^2 \log |\mathcal{A}|.$$

Maintenant on utilise 9.3, "propriétés de l'entropie mesurée", item 3., qui dit que l'entropie est fonction concave de la mesure :

$$\begin{aligned}\sum_{k=0}^{q-1} \sum_{r=0}^{a(k)-1} H_{T_*^{rq+k}\nu_n}(\mathcal{A}_{T,q}) &\leq \sum_{j=0}^{n-1} H_{T_*^j\nu_n}(\mathcal{A}_{T,q}) \\ &\leq nH_{\frac{1}{n}\sum T_*^j\nu_n}(\mathcal{A}_{T,q}) \\ &= nH_{T_*^j\mu_n}(\mathcal{A}_{T,q})\end{aligned}$$

d'où (9.5).

Maintenant, choisissons une suite n_j telle que

$$\frac{1}{n_j} \log |E_{n_j}| \longrightarrow \limsup \frac{1}{n} \log |E_n|$$

et prenons un point d'accumulation μ, pour la topologie faible*, de la suite μ_n. Notons que μ est T-invariante parce que

$$T_*\mu_n - \mu_n = \frac{1}{n}\left(T_*^n\nu_n - \nu_n\right) \longrightarrow 0$$

faiblement*. D'autre part, pour tout $A \in \mathcal{A}_{T,q}$, on a, par définition de la topologie faible*, $\mu_n(A) \longrightarrow \mu(A)$, donc

$$H_{\mu_{n_j}}(\mathcal{A}_{T,q}) \longrightarrow H_\mu(\mathcal{A}_{T,q})$$

d'où, en divisant (9.5) par n,

$$H_\mu(\mathcal{A}_{T,q}) \geq q \limsup \frac{1}{n} \log |E_n|$$

puis, en divisant par q et prenant la limite pour $q \to \infty$,

$$h_\mu(T, \mathcal{A}) \geq \limsup \frac{1}{n} \log |E_n|$$

et enfin, en prenant le sup sur $\mathcal{A}$ (attention tout de même, on a supposé que le diamètre des éléments de $\mathcal{A}$ est $< \epsilon$), on conclut la preuve de la proposition. □

A présent nous pouvons démontrer le théorème 9.4.1.

Proof. Il y a un sens facile avec la Proposition 9.4.2. Soit $\epsilon > 0$. Prenons, pour chaque $n \in \mathbb{N}$, une partie (n, ϵ)-séparée maximale de X. On a (rappelez-vous du cours sur l'entropie topologique)

$$\limsup \frac{1}{n} \log |E_n| = h(T, \epsilon).$$

D'après 9.4.2, il existe une mesure μ_ϵ telle que

$$h(T, \epsilon) = \limsup \frac{1}{n} \log |E_n| \leq h_{\mu_\epsilon}(T).$$

En faisant tendre ϵ vers 0, on obtient une famille de mesures μ_ϵ indicée par ϵ telles que

$$\liminf_{\epsilon \longrightarrow 0} h_{\mu_\epsilon}(T) \geq \lim_{\epsilon \longrightarrow 0} h(T, \epsilon) = h(T)$$

donc $sup_\mu h_\mu(T) \geq h(T)$.

Remarque 9.4.4. *Cela ne prouve pas l'existence d'une mesure d'entropie maximale. L'entropie n'est pas continue comme fonction de la mesure : penser à des mesures portées par des orbites périodiques, donc d'entropie nulle, qui tendent vers une mesure d'entropie* > 0.

Pour l'autre sens, on se donne une mesure μ et on veut montrer que $h_\mu(T) \leq h(T)$. Cela revient à montrer que pour toute partition $\mathcal{A} = \{A_1, \ldots, A_k\}$, on a $h_\mu(T, \mathcal{A}) \leq h(T)$. Pour relier les deux côtés de l'inéquation il faut fabriquer, à partir de la partition $\mathcal{A}$, un ensemble (n, ϵ)-séparé.

On commence par prendre, pour chaque $A_i \in \mathcal{A}$, un compact $B_i \subset A_i$, tel que la partition $\mathcal{B} = \{B_0, B_1, \ldots, B_k\}$, où $B_0 = X \setminus \cup_{i=1}^k B_i$, vérifie $H(\mathcal{A}|\mathcal{B}) < 1$. C'est possible car il suffit pour cela que

$$\sum_{j=0}^{k} \mu(B_j) \sum_{i=1}^{k} \mu(A_i|B_j) \log \mu(A_i|B_j) < 1,$$

or

$$\begin{aligned} \mu(A_i|B_j) =& 0 \text{ si } j \neq 0 \text{ et } j \neq i \\ & 1 \text{ si } j \neq 0 \text{ et } j = i \end{aligned}$$

donc

$$H(\mathcal{A}|\mathcal{B}) = \sum_{i=1}^{k} \mu(A_i \cap B_0) \log \frac{\mu(A_i \cap B_0)}{\mu(B_0)}$$

et, puisque μ est borélienne, on peut choisir $\mu(B_i)$ aussi proche qu'on veut de $\mu(A_i)$, du coup $\mu(A_i \cap B_0)$ est arbitrairement proche de 0. Alors (9.3, item 1.), on a $h_\mu(T, \mathcal{A}) \leq h_\mu(T, \mathcal{B}) + 1$.

L'intérêt d'avoir remplacé $\mathcal{A}$ par $\mathcal{B}$, c'est que les compacts sont plus sympathiques que les mesurables. En particulier, maintenant on a les recouvrements ouverts

$$\mathcal{U} = \{B_0 \cup B_1, \ldots, B_0 \cup B_k\}$$

dont les éléments sont bien des ouverts car, par exemple,

$$B_0 \cup B_1 = X \setminus \bigcup_{i=2}^{k} B_i$$

et (de la même façon qu'on construit $\mathcal{A}_{T,n}$ à partir de $\mathcal{A}$)

$$\mathcal{U}_{T,n} = \left\{ \bigcap_{i=0}^{n-1} T^{-i} U_j / U_j \in \mathcal{U} \right\}.$$

Prenons un sous-recouvrement minimal (i.e. si on enlève un élément, ce n'est plus un recouvrement) de $\mathcal{U}_{T,n}$, notons-le $\mathcal{U}'_n$. Observons que

$$T^{-1}(B_0 \cup B_j) = T^{-1}(B_0) \cup T^{-1}(B_j),$$

donc (par récurrence sur n), tout élément de $\mathcal{U}_{T,n}$ est réunion d'au plus 2^n éléments de $\mathcal{B}_{T,n}$. De plus, les éléments de $\mathcal{B}_{T,n}$ étant disjoints, chacun d'entre eux est contenu dans un élément de $\mathcal{U}'_n$, puisque $\mathcal{U}'_n$ recouvre X. Par conséquent

$$2^n |\mathcal{U}'_n| \geq |\mathcal{B}_{T,n}|.$$

Maintenant, soit δ_0 le supremum des $r > 0$ tels que toute boule de rayon r est contenue dans un élément de $\mathcal{U}$ (c'est le nombre de Lebesgue du recouvrement $\mathcal{U}$ pour la distance d), $\delta_0 > 0$ par compacité de X.

Alors (cf. preuve de 9.4.3) δ_0 est aussi le nombre de Lebesgue du recouvrement $\mathcal{U}_{T,n}$ pour la distance d_n^T. Donc, pour tout élément U de $\mathcal{U}'_n$, il existe x_U, tel que les $x_U, U \in \mathcal{U}'_n$ forment une partie (n, δ_0)-séparée de X. Donc

$$\begin{aligned} h(T) &\geq \limsup \frac{1}{n} \log |\mathcal{U}'_n| \\ &\geq \limsup \frac{1}{n} \log |\mathcal{B}_{T,n}| - \log 2. \end{aligned}$$

Maintenant on se souvient de la majoration de l'entropie d'une partition : $H(\mathcal{A}) \leq \log |\mathcal{A}|$. On a donc

$$\begin{aligned} h(T) &\geq \limsup \frac{1}{n} H(\mathcal{B}_{T,n}) - \log 2 \\ &= h_\mu(T, \mathcal{B}) - \log 2 \\ &\geq h_\mu(T, \mathcal{A}) - \log 2 - 1. \end{aligned}$$

Cela ne fait pas tout à fait notre affaire, mais là on se souvient que $h(T^k) = kh(T)$ et $h_\mu(T^k) = kh_\mu(T)$ (9.3, "propriétés de l'entropie mesurée", item 2.). On peut donc remplacer T par T^K dans le calcul qui précède, et en faisant tendre k vers l'infini, on obtient $h(T) \geq h_\mu(T)$. $\square$

Chapter 10

Entropie volumique

Ce chapitre est moins élémentaire que les précédents et nécessite des notions de base de topologie algébrique et de géométrie riemannienne.

Dans ce chapitre, X est le fibré unitaire tangent d'une variété riemannienne (M, g), et T_t est le flot géodésique. En d'autres termes, X est l'ensemble des couples (position initiale dans X, direction dans laquelle on veut aller), et le flot géodésique consiste à partir tout droit (i.e. en suivant une géodésique, c'est un dire un plus court chemin) dans la direction choisi, à vitesse $1m/s$ (il y a des mètres puisque notre variété est riemannienne, on peut mesurer des longueurs, pour ce qui est des secondes...), pendant un temps t. On notera $\tilde{M}$ le revêtement universel de M. On note encore g la métrique relevée à $\tilde{M}$. Le groupe fondamental agit alors par isométries sur $\tilde{M}$. Pour $x \in \tilde{M}$, on note $B(x, r)$ la boule de centre x et de rayon r dans $\tilde{M}$, et on note $V(x, r)$ le volume de $B(x, r)$.

10.1 Définition

Définition 10.1.1. *On appelle entropie volumique de* (M, g)*, et on note* $h_{vol}(M, g)$*, la limite suivante :*

$$\lim_{r\to\infty} \frac{1}{r} \log V(x, r).$$

Preuve de l'existence de la limite.

Soit N un domaine fondamental de l'action sur $\tilde{M}$ du groupe fondamental de M, et soit d son diamètre (pour g). Commençons par observer que

$$\forall r \geq d, \forall x, y \in N,\ B(x, r-d) \subset B(y, r) \subset B(x, r+d). \tag{10.1}$$

En effet, si $z \in B(x, r-d)$, on a $d(x, z) \leq r - d$, et puisque $x, y \in N$, qui est de diamètre d, on a $d(x, y) \leq d$, donc par l'inégalité triangulaire, $d(y, z) \leq$

$d(x,y)+d(x,z) \leq r$, donc $z \in B(y,r)$. Ceci montre que $B(x,r-d) \subset B(y,r)$. L'autre inclusion est un exercice pour vous.

En prenant les volumes dans (10.1), on obtient

$$\forall r \geq d, \forall x,y \in N,\ V(x,r-d) \leq V(y,r) \leq V(x,r+d). \tag{10.2}$$

Mais pour tous $x,y \in \tilde{M}$, il existe des isométries (pas forcément les mêmes pour x et pour y !) qui envoient x et y dans N, et les isométries conservent le volume, donc

$$\forall r \geq d, \forall x,y \in \tilde{M},\ V(x,r-d) \leq V(y,r) \leq V(x,r+d). \tag{10.3}$$

Maintenant, prenons $\delta > 0$, et prenons une partie δ-séparée maximale Y de $B(x,r)$ (δ-séparée signifie que deux points distincts de Y sont distants d'au moins δ ; maximale s'entend au sens de l'inclusion. L'existence, et la finitude, de Y résulte de la compacité de $B(x,r)$).

Puisque Y est δ-séparée, les boules de rayon $\delta/2$ centrées en les points de Y sont disjointes, et par l'inégalité triangulaire, elles sont toutes contenues dans $B(x,r+\delta/2)$. Le volume de $B(x,r+\delta/2)$ est donc au moins égal à la somme des volumes des boules de rayon $\delta/2$ centrées en les points de Y. Soit

$$C_\delta = \min_{y \in Y} V(y,\delta/2).$$

On a alors

$$V(x,r+\delta/2) \geq \#(Y) C_\delta \tag{10.4}$$

où $\#(Y)$ est le cardinal de Y.

Mais, puisque Y est maximale, tout point de $B(x,r)$ est à moins de δ d'un point de Y, sans quoi on pourrait l'ajouter à Y, ce qui contredit la maximalité de Y.

Maintenant, prenons $s > 0$. Tout point de $B(x,r+s)$ est à distance $\leq s$ d'un point de $B(x,r)$ (cette propriété n'a rien d'évident et résulte de l'existence d'une géodésique entre x et le point considéré), donc, par l'observation ci-dessus, à distance $\leq s+\delta$ d'un point de Y. On a donc

$$B(x,r+s) \subset \bigcup_{y \in Y} B(y,s+\delta)$$

et par conséquent

$$V(x,r+s) \leq \sum_{y \in Y} V(y,s+\delta)$$

d'où, avec (10.2) (2ème inégalité)

$$V(x,r+s) \leq \sum_{y \in Y} V(x,s+\delta+d) = \#(Y) V(x,s+\delta+d).$$

En appliquant (10.4), on obtient

$$V(x, r+s) \leq \frac{V(x, r+\delta/2)}{C_\delta} V(x, s+\delta+d).$$

On conclut ensuite en utilisant le lemme (désormais) bien connu sur les fonctions sous-additives (exercice 82). □

10.2 Théorème de Manning

Jusqu'ici vous vous demandez peut-être ce que cette entropie volumique a à voir avec une entropie. Voici la réponse.

Théorème 10.2.1 (Manning, Annals of Mathematics, 1978)**.** *En notant $h_{top}(M, g)$ l'entropie topologique du flot géodésique (flot sur le fibré unitaire tangent, qui est compact), on a $h_{top}(M, g) \geq h_{vol}(M, g)$.*

Ce théorème peut faire l'objet d'un exposé. Un autre exposé possible concerne le calcul exact de l'entropie du flot géodésique d'une surface hyperbolique, ainsi que la détermination de la mesure d'entropie maximale. On commence par montrer l'inégalité. On veut montrer que pour tout $\epsilon > 0$, on a $h_{top}(M, g) \geq h_{vol}(M, g) - \epsilon$. Soit donc $\epsilon > 0$. Par définition de l'entropie volumique, il existe $r(\epsilon)$ tel que $\forall r \geq r(\epsilon)$,

$$\exp(h_{vol}(M, g) + \epsilon)r \geq V(x, r) \geq \exp(h_{vol}(M, g) - \epsilon)r.$$

Maintenant prenons $\delta > 0$ (on ajustera la valeur plus tard) et montrons l'existence d'une suite $(r_k)_{k\in\mathbb{N}}$ telle que $\lim_{k\to\infty} r_k = +\infty$, et

$$\forall k \in \mathbb{N},\ V(x, r_k + \frac{\delta}{2}) - V(x, r_k) \geq \exp\left((h_{vol}(M, g) - \epsilon)r_k\right). \tag{10.5}$$

En effet, considérons la suite

$$V_n = V(x, (n+1)\frac{\delta}{2}) - V(x, n\frac{\delta}{2}).$$

On a

$$\sum_{n=0}^{N} V_n = V(x, (N+1)\frac{\delta}{2}) \geq \exp\left((h_{vol}(M, g) - \epsilon)(N+1)\frac{\delta}{2}\right)$$

pour N assez grand, mais s'il existait N_0 tel que

$$\forall n \geq N_0,\ V_n \leq \exp\left((h_{vol}(M, g) - \epsilon)n\frac{\delta}{2}\right)$$

alors (rappelez-vous des comparaisons série-intégrale en L2), la somme de la série de terme général V_n ne pourrait pas être équivalente à $\exp(h_{vol}(M,g)n)$, donc il existe une suite d'entiers naturels $n_k \to \infty$ telle que

$$\forall k \in \mathbb{N}, V_{n_k} \geq \exp\left((h_{vol}(M,g) - \epsilon)n_k \frac{\delta}{2}\right),$$

ce qui prouve (10.5) avec $r_k = n_k \frac{\delta}{2}$.

Maintenant prenons, pour chaque $k \in \mathbb{N}$, une partie 2δ-séparée maximale Q_k de $B(x, r_{k+1}) \backslash B(x, r_k)$. Alors les boules de rayon 2δ centrées en les points de Q_k recouvrent $B(x, r_{k+1}) \setminus B(x, r_k)$, sans quoi Q_k ne serait pas maximale (au sens de l'inclusion, parmi les parties 2δ-séparées), donc

$$V(x, r_{k+1}) - V(x, r_k) \leq |Q_k| \sup_{z \in M} V(z, 2\delta).$$

On se rapproche un peu du but, qui est de minorer l'entropie topologique, qui est (la limite quand δ tend vers zéro du) taux de croissance exponentielle d'une partie (t, δ)-séparée de X (rappelez-vous que X est le fibré unitaire tangent à M) : on a montré que le cardinal d'une partie 2δ-séparée d'une grande couronne circulaire dans $\tilde{M}$ croît exponentiellement avec le rayon de la couronne.

Pour autant on n'est pas tiré d'affaire : ce qu'on veut, c'est une partie $(t, 2\delta)$-séparée de X, et ce qu'on a, c'est une partie 2δ-séparée de $\tilde{M}$. Il y a trois boulons à resserrer :

- (t, δ)-séparée est différent de δ-séparée, on ne mesure pas les distances de la même façon
- X est formé de couples (position, vitesse), alors que dans $\tilde{M}$ il n'y a que des positions
- une position dans $\tilde{M}$, ce n'est pas la même chose qu'une position dans M : il y a en général une infinité de positions dans $\tilde{M}$ qui correspondent à une même position dans M, il peut être difficile de comparer le volume d'une boule dans $\tilde{M}$ avec le volume de la boule correspondante dans M.

Le premier boulon est la raison pour laquelle on a pris une couronne circulaire $B(x, r_{k+1}) \setminus B(x, r_k)$.

10.3 Courbure négative

Théorème 10.3.1 (Manning, Annals of Mathematics, 1978). *Sous les hypothèses du théorème 10.2.1, et si de plus (M, g) est à courbure sectionnelle négative, on a l'égalité $h_{top}(M, g) = h_{vol}(M, g)$.*

Contents