# Systèmes dynamiques I

Patrice Le Calvez

Cours fondamental I, Master 2

2014-2015

## 0.1 Origine des systèmes dynamiques.

On fait souvent remonter l'origine des systèmes dynamiques à l'astronomie et plus particulièrement à la mécanique céleste. Les lois de la gravitation de Newton nous disent que le déplacement de $n$ corps dans l'espace $\mathbf{R}^3$, de masses respectives $m_1$, $m_2$, $\ldots$, $m_n$, est régit par l'équation différentielle suivante

$$m_i \frac{d^2\vec{x}_i}{dt^2} = -G \sum_{j \neq i} m_i m_j \frac{\vec{x}_i - \vec{x}_j}{\|\vec{x}_i - \vec{x}_j\|^3},$$

où $\vec{x}_i \in \mathbf{R}^3$ est la position du $i$-ème corps et $G$ la constante de gravitation. Dans le cas où $n = 2$, on peut intégrer le système : le centre de gravité $M$ de ce système décrit un mouvement rectiligne uniforme et les deux corps décrivent chacun une conique de foyer $M$ dans le repère galiléen dont l'origine est en $M$.

Dans le cas du système solaire, chaque planète décrit approximativement une ellipse. L'approximation est due à l'attraction (faible) des autres planètes du système. Cependant, on ne peut pas décrire le mouvement exact des planètes à cause du fait fondamental suivant : il est impossible d'intégrer le système d'équations différentielles écrit plus haut dès que $n \geq 3$. Un des buts des systèmes dynamiques est d'expliquer pourquoi certains systèmes sont intégrables et d'autres non, et de décrire la complexité de ceux qui ne sont pas intégrables.

On peut, à l'aide de méthodes numériques, donner des solutions approchées et décrire l'évolution du système sur des grands intervalles du temps. C'est ainsi par exemple, qu'à l'aide à la méthode des perturbations, John Adams et Urbain Le Verrier ont découvert l'existence de Neptune. Cependant, ces méthodes ne permettent pas de répondre à des questions structurelles sur le système comme, par exemple, sur la stabilité du système solaire : la terre peut-elle être éjectée du système solaire ? peut-elle s'écraser sur le soleil ? Ce type de questions sont appelées "qualitatives", on ne cherche pas à calculer les solutions, mais à comprendre leur comportement. Voici d'autres types de questions qualitatives :

- que se passe t-il quand $t$ tend vers $+\infty$ ?
- existe t-il des solutions pérodiques ? sont-elles "nombreuses" ?
- si on perturbe les conditions initiales du système, le comportement du mouvement reste-t-il proche de celui du mouvement initial ou au contraire est-il radicalement différent ?

Cette idée de s'intéresser aux questions qualitatives, plus précisément à la géométrie de l'espace des solutions, est due à Henri Poincaré, l'un des pères des systèmes dynamiques.

Dans le cas du problème à $n$ corps, il y a beaucoup de variables réelles (un peu moins que $6 \times n$ à cause des symétries). On pourrait penser que c'est le grand nombre de variables qui est cause de la non-intégrabilité du système. En fait, beaucoup de systèmes, avec bien moins de variables, partagent cette propriété. Il en est ainsi par exemple du pendule double, où un corps est attaché à un autre par une tige rigide, lui même attaché à une origine fixe par une autre tige rigide. Il n'y a pourtant que quatre variables. Tous les exemples précédents sont issus de la mécanique et appartiennent de ce fait à la classe des *systèmes conservatifs*, où certaines quantités importantes sont conservées (ici l'énergie totale). L'imprédicabilité se retrouve également dans des systèmes dissipatifs. Par exemple, dès 1963, le météorologue Edward Lorenz a découvert que le système différentiel dissipatif suivant

$$\begin{cases} x' = 10(y-x) \\ y' = x(28-z) - y \\ z' = xy - \frac{8}{3}z \end{cases}$$

présentait une forte dépendance par rapport aux conditions initiales.

Les systèmes dynamiques ne se restreignent pas à l'étude des équations différentielles et incluent également d'autres types de systèmes d'évolution. En voici un exemple : considérons une application $f : I \to I$ définie sur un intervalle $I$, fixons un point $u_0 \in I$ et étudions la suite $(u_n)_{n\geq 0}$ définie par la relation de récurrence suivante $u_{n+1} = f(u_n)$. Dans beaucoup de cas (en particulier dans les cas étudiés dans les premières années d'université) cette suite converge vers un point fixe de $f$ . Cependant le comportement peut être bien plus riche, même quand les formules de récurrence sont simples. Un exemple classique est *l'application logistique*

$$f_\lambda \; : [0,1] \to [0,1] \atop x \mapsto \lambda x(1-x),$$

où $\lambda \in [0,4]$. Si on choisit $u_0 = 0$, la suite est constante, mais si on choisit $u_0 = 1/2$, le comportement de la suite sera très différent suivant les valeurs de $\lambda$. Dans certains cas, la suite s'approchera d'une suite périodique, dans d'autres cas, elle prendra des valeurs denses dans l'intervalle, avec un comportement bien plus complexe.

Un des buts des systèmes dynamiques est d'introduire un langage commun qui puisse intégrer tous ces types de systèmes d'évolutions déterministes. Une première étape (que l'on initiera ici) est d'introduire un certain nombre de modèles, où apparaît déjà une complexité, mais où cette complexité peut-être expliquée très précisément.

Pour finir cette section, on peut noter que des problèmes de mathématiques, a priori non liés à des systèmes d'évolution ont des explications dynamiques. C'est le cas, par exemple en théorie des nombres de problèmes hautement non triviaux. Donnons ici un exemple (qui lui est très simple !) : considérons les puissances de 2 et ne gardons que la première unité, on trouve : 1, 2, 4, 8, 1, 3, 6, 1, 2, 5,.... Quelle est la proportion de 7 dans cette suite ?

## 0.2 Quelques rappels sur les équations différentielles ordinaires.

À un champ de vecteurs $F \; : U \to \mathbf{R}^n$ défini sur une partie ouverte $U$ de $\mathbf{R}^n$ (ou plus généralement à un champ de vecteurs défini sur une variété lisse) est naturellement associée une équation différentielle $x' = F(x)$. Une *solution* de cette équation est un couple $(I, \varphi)$, où $I$ est un intervalle de $\mathbf{R}$ et $\varphi \; : I \to U$ une fonction dérivable qui vérifie $\varphi'(t) = F(\varphi(t))$ pour tout $t \in I$. Elle est *maximale* s'il n'existe aucune solution $(J, \psi)$ telle que $J$ contienne strictement $I$ et telle que $\psi|_I = \varphi$. Le résultat qui suit est une conséquence directe du théorème de Cauchy-Lipschitz sur l'existence et l'unicité locale des solutions.

Théorème 0.2.1 : *Si $F$ est de classe $C^1$, alors pour tout $t_0 \in \mathbf{R}$ et tout $x_0 \in U$, il existe une unique solution maximale $(I, \varphi)$ telle que $t_0 \in I$ et $\varphi(t_0) = x_0$.*

**Exemples**

1. Supposons que $U = \mathbf{R}^n$ et que $F = A$ est une application linéaire. Alors $I = \mathbf{R}$ et $\varphi(t) = e^{(t-t_0)A}(x_0)$, où $e^{(t-t_0)A}$ est l'exponentielle de l'application linéaire $(t - t_0)A$.

2. Supposons que $U = \mathbf{R}$ et que $F(x) = x^2$. Alors

- si $x_0 = 0$, on a $I = \mathbf{R}$ et $\varphi(t) = 0$ ;
- si $x_0 > 0$, on a $I = ]-\infty, t_0 + \dfrac{1}{x_0}[$ et $\varphi(t) = \dfrac{x_0}{1 + x_0(t_0 - t)}$ ;
- si $x_0 < 0$, on a $I = ]t_0 + \dfrac{1}{x_0}, +\infty[$ et $\varphi(t) = \dfrac{x_0}{1 + x_0(t_0 - t)}$.

**Remarques**

1. Le système étant autonome, on a $I = J + t_0$ et $\varphi(t) = \psi(t - t_0)$, où $(J, \psi)$ est la solution maximale de condition initiale $\psi(0) = x_0$.

2. Si $F$ est de classe $C^r$, alors $\varphi$ est de classe $C^{r+1}$.

Un champ de vecteurs est *complet* si toute solution maximale est définie sur $\mathbf{R}$. C'est le cas du premier exemple précédent, mais pas du second. Une condition suffisante pour qu'un champ de vecteurs sur $\mathbf{R}^n$ soit complet est l'égalité $\|F(x)\| \leq A\|x\| + B$ pour tout $x \in \mathbf{R}^n$, où $\| \ \|$ est une norme et $A$, $B$ des constantes réelles. Une condition suffisante pour qu'un champ de vecteurs sur une variété soit complet est la compacité de la variété. Dans le cas d'un champ de vecteurs complet, on peut définir une application

$$\begin{aligned} \Phi \ &: U \times \mathbf{R} \to U \\ &(x, t) \mapsto \Phi(x, t) = \varphi_x(t) \end{aligned},$$

où $(\mathbf{R}, \varphi_x)$ est la solution maximale de condition initiale $\varphi_x(0) = x$. Le théorème 1.2.1 peut alors être amélioré

Théorème 0.2.2 : *Si $F$ est de classe $C^r$, l'application $\Phi$ est de classe $C^r$.*

Pour tout $t \in \mathbf{R}$ on peut définir une application

$$\begin{aligned} \theta_t \ &: U \to U \\ &x \mapsto \theta_t(x) = \Phi(x, t) \end{aligned}.$$

On obtient :

Théorème 0.2.3 : **i)** *Chaque $\theta_t$ est un diffféomorphisme de classe $C^r$ de $U$ ;*
**ii)** *on a $\theta_0 = \mathrm{Id}_U$ ;*
**iii)** *on a $\theta_{t+t'} = \theta_t \circ \theta_{t'}$ pour tous réels $t$ et $t'$.*

*Preuve.* L'assertion **ii)** provient de l'égalité $\theta_0(x) = \varphi_x(0) = x$. Pour montrer **iii)**, remarquons que $\theta_t \circ \theta_{t'}(x) = \varphi_{\varphi_x(t')}(t)$. L'application $\varphi_{\varphi_x(t')}$ est la solution maximale de condition initiale $\varphi_{\varphi_x(t')}(0) = \varphi_x(t')$. L'équation $x' = F(x)$ étant autonome, on en déduit que

$$\varphi_{\varphi_x(t')}(t) = \varphi_x(t + t') = \theta_{t+t'}(x).$$

Le théorème 0.2.2 nous dit que chaque $\theta_t$ est de classe $C^r$. Pour montrer que c'est un difféomorphisme, il suffit de montrer qu'elle est inversible et que $(\theta_t)^{-1} = \theta_{-t}$. Or on a

$$\theta_t \circ \theta_{-t} = \theta_{t-t} = \theta_0 = \mathrm{Id}_U = \theta_{-t} \circ \theta_t.$$

$\square$

Le résultat précédent peut être résumé en disant que l'application $t \mapsto \theta_t$ est un morphisme du groupe additif $(\mathbf{R}, +)$ dans le groupe $(\mathrm{Diff}^r(U), \circ)$ des difféomorphismes de classe $C^r$ de $U$, muni de la loi de composition. On dit que la famille $(\theta_t)_{t\in\mathbf{R}}$ est un *groupe à un paramètre de difféomorphismes*, ou encore un *flot*. Un tel processus déterministe définit un système dynamique.

## 0.3 Qu'est ce qu'un système dynamique ?

En toute généralité un système dynamique est une action d'un groupe, ou plus généralement d'un semi-groupe, $G$ sur un ensemble $X$, c'est-à-dire une famille d'applications $\theta_g\ : X \to X$, telle que $\theta_e = \mathrm{Id}_X$ et $\theta_{gg'} = \theta_g \circ \theta_{g'}$, pour tous $g$ et $g'$ dans $G$. Comme on est intéressé par les propriétés asymptotiques de l'action, on se limite au cas des groupes topologiques non compacts et plus particulièrement aux semi-groupes $\mathbf{N}$, $\mathbf{Z}$, $[0, +\infty[$, $\mathbf{R}$. Dans les deux premiers cas on a un *système dynamique discret*, dans les deux autres, un *système dynamique continu*. Pour une action de $\mathbf{N}$, on a $\theta_n = \theta_1^n$. Un tel système est donc défini par les itérés de l'application $T = \theta_1\ : X \to X$. Une action de $\mathbf{Z}$ est définie par les itérés positifs et négatifs d'une bijection $T\ : X \to X$. Dans la section précédente, nous avons donné des exemples d'actions naturelles de $\mathbf{R}$ sur des variétés. Notons que l'on a une action de $[0, +\infty[$ chaque fois que l'on a un champ de vecteurs positivement complet, c'est-à-dire pour lequel les solutions maximales sont définies jusqu'à $+\infty$.

La définition est bien sûr trop générale pour être intéressante et il nous faudra rajouter une structure sur $X$ (et sur $G$ si $G \neq \mathbf{N}$ ou $\mathbf{Z}$). Par exemple, si $X$ est un espace topologique, $G$ un groupe topologique et si chaque $(g, x) \mapsto \theta_g(x)$ est continue, on obtient un *système dynamique topologique*. Si $X$ est une partie ouverte de $\mathbf{R}^n$ ou plus généralement une variété lisse, $G$ un groupe de Lie et si $(g, x) \mapsto \theta_g(x)$ est différentiable, on a un *système dynamique différentiable*. Un autre cas intéressant est le cas où $X$ est muni d'une $\sigma$-algèbre $\mathcal{B}$ et où chaque $T_g$ est mesurable : on a alors un *système dynamique mesurable*. Si, de plus, il existe une mesure $\mu$ sur $\mathcal{B}$ qui est préservée par chaque $T_g$, c'est-à-dire si pour tout $A \in \mathcal{B}$ et tout $g \in G$, on a $\mu(T_g^{-1}(A)) = \mu(A)$, on a un *système dynamique mesuré*. L'étude des systèmes dynamiques mesurés est appelée *théorie ergodique*.

Les flots introduits dans la section précédente appartiennent à la catégorie des systèmes dynamiques différentiables mais certaines de leurs propriétés sont obtenues en ne considérant que le caractère topologique du système. La théorie ergodique peut également être nécessaire à l'étude de certains flots, comme, par exemple, les *systèmes hamiltoniens* (qui proviennent de la mécanique comme dans la première section). Pour définir de tels flots, on se donne une fonction $H\ : U \to \mathbf{R}$ de classe $C^2$ sur une partie ouverte de $\mathbf{R}^{2n}$ (ou plus généralement sur une *variété symplectique*). Un système hamiltonien sur $U$ est défini par les équations

$$\begin{cases} x_1' = \frac{\partial H}{\partial y_1}(x_1, \ldots, x_n, y_1, \ldots, y_n) \\ \ldots = \ldots \\ x_n' = \frac{\partial H}{\partial y_n}(x_n, \ldots, x_n, y_1, \ldots, y_n) \\ y_1' = -\frac{\partial H}{\partial x_1}(x_1, \ldots, x_n, y_1, \ldots, y_n) \\ \ldots = \ldots \\ y_n' = -\frac{\partial H}{\partial x_n}(x_1, \ldots, x_n, y_1, \ldots, y_n) \end{cases}.$$

Si le champ de vecteurs

$$F = (\frac{\partial H}{\partial y_1}, \ldots, \frac{\partial H}{\partial y_n}, -\frac{\partial H}{\partial x_1}, \ldots, -\frac{\partial H}{\partial x_n})$$

est complet, le flot induit $(\theta_t)_{t\in\mathbf{R}}$ définit un système dynamique mesuré. En effet, chaque $\theta_t$ préserve la forme différentielle

$$\omega = \sum_{i=1}^{n} dx_i \wedge dy_i,$$

et par conséquent préserve également la forme volume

$$\omega \wedge \ldots \wedge \omega = dx_1 \wedge \ldots \wedge dx_n \wedge dy_1 \wedge \ldots \wedge dy_n.$$

Il en est ainsi de l'équation du pendule libre $x'' + l \sin x = 0$, que l'on peut écrire

$$\begin{cases} x' = y \\ y' = -l \sin x \end{cases}.$$

C'est un système hamiltonien complet dont l'hamiltonien

$$H \ : (x, y) \mapsto \frac{1}{2}y^2 - l \cos x$$

est l'énergie totale. Ici, le flot préserve la forme d'aire $dx \wedge dy$.

On s'intéressera particulièrement aux systèmes dynamiques discrets. Bien souvent, un système continu $(\theta_t)_{t\in\mathbf{R}}$ peut se ramener à un système discret. Une façon naïve est de ne s'intéresser qu'aux valeurs entières de $t$, c'est-à-dire d'étudier les itérés de $\theta_1$. On perd cependant beaucoup d'informations en agissant ainsi. Une réduction n'a généralement d'intérêt que si l'on obtient un système discret sur un espace de dimension plus petite. On va donner un exemple. Donnons-nous un champ de vecteurs $F$ dépendant du temps sur une partie ouverte $U \subset \mathbf{R}^n$ mais périodique en temps $(F(x, t+T) = F(x, t))$.Comme dans le cas des systèmes autonomes, on obtient une famille $(\theta_t)_{t\in\mathbf{R}}$ de difféomorphismes de $U$, mais qui n'est plus un flot. On a cependant $\theta_{t+T} = \theta_t \circ \theta_T$, pour tout $t \in \mathbf{R}$. Le système peut se réduire à l'étude des itérés de $\theta_T$. Considérons par exemple le pendule entretenu périodique d'équation $x'' + l \sin x = p(t)$, où la fonction $p$ est périodique. C'est un système autonome sur $\mathbf{R}^3$ :

$$\begin{cases} x' = y \\ y' = -l \sin x + p(\tau) \\ \tau' = 1 \end{cases},$$

que l'on peut réduire à un système discret sur $\mathbf{R}^2$. Concluons cette section en indiquant que la dynamique du pendule simple est facile à décrire, et sans forte dépendance par rapport aux conditions initiales, et que ce n'est plus du tout le cas dès que l'on ajoute un terme perturbatif périodique.

## 1.1 Orbites, différents types de récurrence.

Un système dynamique topologique (discret) est défini par une application continue $T : X \to X$, où $X$ est un espace topologique. On veut étudier l'action naturelle de $\mathbf{N}$ obtenue par les itérés $T^n = T \circ \ldots \circ T$ ($n$ fois), et plus particulièrement par les propriétés asymptotiques de $T^n$. Si $T$ est un homéomorphisme, on peut définir $T^k$, $k \leq 0$, en posant $T^k = (T^{-k})^{-1}$, obtenant ainsi une action de $\mathbf{Z}$.

**Définition.** L'*orbite positive* d'un point $x \in X$ est l'ensemble $O^+(x) = \{T^n(x)\,, n \geq 0\}$. Si $T$ est un homéomorphisme, on peut définir l'*orbite négative* $O^-(x) = \{T^n(x)\,, n \leq 0\}$ et l'*orbite totale* ou *orbite* $O(x) = O^-(x) \cup O^+(x) = \{T^k(x)\,, k \in \mathbf{Z}\}$.

**Exemple.** Pour l'application

$$\begin{aligned} T\ : \mathbf{Z} &\to \mathbf{Z} \\ x &\mapsto x+1 \end{aligned}$$

on a $O^+(0) = \mathbf{N}$, $O^-(0) = -\mathbf{N}$ et $O(0) = \mathbf{Z}$.

**Remarque.** Dans le cas non inversible, on parle souvent d'orbite au lieu d'orbite positive, en l'absence d'ambiguité.

**Définition.** Un point $x \in X$ est un *point périodique* de $T$ s'il existe $k \geq 1$ tel que $T^k(x) = x$. Le plus petit entier strictement positif vérifiant cette propriété est la *période* de $x$.

PROPOSITION 1.1.1 : *Soit $x$ un point périodique de $T : X \to X$, de période $q$. Alors :*

**i)** *$O^+(x)$ a $q$ éléments, les points $x$, $T(x)$, $\ldots$ et $T^{q-1}(x)$ ;*

**ii)** *pour tout $n \geq 0$, on a $T^n(x) = T^r(x)$, où $r$ est le reste de la division euclidienne de $n$ par $q$ ;*

**iii)** *si $T$ est un homéomorphisme, on a $O(x) = O^+(x)$ et pour tout $k \in \mathbf{Z}$, on a $T^k(x) = T^r(x)$, où $r$ est le reste de la division euclidienne de $k$ par $q$.*

*Preuve.* Par une simple récurrence sur $s$, on sait que $T^{qs}(x) = x$ pour tout $s \geq 0$. Si $n = qs+r$, $0 \leq r < q$, est la division euclidienne de $n$ par $q$, on a $T^n(s) = T^r(T^{qs}(x)) = T^r(x)$. L'assertion **ii)** est donc prouvée. Pour obtenir **i)**, il reste à montrer que les points $x$, $T(x)$, $\ldots$, $T^{q-1}(x)$ sont distincts. Si ce n'est pas le cas, alors il existe $n$, $n'$, tels que $T^n(x) = T^{n'}(x)$ et $0 \leq n < n' < q$. On en déduit

$$T^{n'-n}(x) = T^{n'-n}(T^q(x)) = T^{n'-n+q}(x) = T^{q-n}(T^{n'}(x)) = T^{q-n}(T^n(x)) = T^q(x) = x,$$

ce qui contredit la définition de $q$, puisque $1 \leq n'-n < q$. Pour établir **iii)**, remarquons que pour tout $s < 0$, on a

$$T^{-qs}(x) = T^{-qs}(T^{qs}(x)) = T^{-qs+qs}(x) = x.$$

Ceci implique que $T^k(x) = T^r(x)$, où $r$ est le reste de la division euclidienne de $k$ par $q$ et donc que $O(x) = O^+(x)$. $\square$

La périodicité est la plus forte propriété de récurrence, nous allons introduire dans cette section des notions de récurrence plus faibles.

**Définition.** Une partie $Y$ de $X$ est *positivement invariante* si $T(Y) \subset Y$. Nous avons alors un système dynamique discret restreint $T|_Y \ : Y \to Y$. Si $T$ est un homéomorphisme, nous dirons que $Y$ est *négativement invariante* si $Y \subset T(Y)$ et *globalement invariante* (ou *invariante*) si $T(Y) = Y$.

**Exemple.** L'ensemble $\mathbf{N}$ est une partie positivement invariante de

$$T \ : \begin{array}{l} \mathbf{Z} \to \mathbf{Z} \\ x \mapsto x+1 \end{array}.$$

**Remarques**

1. Une orbite positive est positivement invariante.

2. Si $Y$ est une partie positivement invariante, il en est de même de son adhérence $\overline{Y}$. En effet, si $x \in \overline{Y}$, alors pour tout voisinage $U$ de $T(x)$, la partie $T^{-1}(U)$ est un voisinage de $x$ (puisque $T$ est continue), ce qui implique que $T^{-1}(U) \cap Y \neq \emptyset$ (puisque $x \in \overline{Y}$). Ainsi, on a

$$\emptyset \neq T(T^{-1}(U) \cap Y) \subset U \cap T(Y) \subset U \cap Y,$$

ce qui implique que $T(x) \in \overline{Y}$.

3. L'union ou l'intersection de parties positivement invariantes $(Y_i)_{i\in I}$ est positivement invariante. En effet, :

$$T(\bigcap_{i\in I} Y_i) \subset \bigcap_{i\in I} T(Y_i) \subset \bigcap_{i\in I} Y_i \ \text{ et } \ T(\bigcup_{i\in I} Y_i) = \bigcup_{i\in I} T(Y_i) \subset \bigcup_{i\in I} Y_i.$$

4. Dans le cas où $T$ est un homéomorphisme, on a des résultats similaires pour les parties négativement ou globalement invariantes.

**Définition.** L'*ensemble $\omega$-limite* d'un point $x \in X$ est l'ensemble, noté $\omega(x)$, des points d'accumulation de la suite $(T^n(x))_{n\geq 0}$ : un point $y$ appartient à $\omega(x)$ s'il existe une suite $(n_i)_{\geq 0}$ d'entiers positifs telle que $\lim_{i\to+\infty} n_i = +\infty$ et $\lim_{i\to+\infty} T^{n_i}(x) = y$.

**Exemples.**

1. Si $x$ est un point périodique, alors $\omega(x) = O^+(x)$.

2. On a $\omega(0) = \emptyset$ si

$$T \ : \begin{array}{l} \mathbf{Z} \to \mathbf{Z} \\ x \mapsto x+1 \end{array}.$$

PROPOSITION 1.1.2 : *L'ensemble $\omega(x)$ est fermé et positivement invariant. Si $T$ est un homéomorphisme, il est globalement invariant.*

*Preuve.* On peut écrire

$$\begin{aligned} \omega(x) &= \bigcap_{m\geq 0} \overline{\{T^n(x)\,,\, n \geq m\}} \\ &= \bigcap_{m\geq 0} \overline{O^+(T^m(x))}, \end{aligned}$$

comme intersection de parties fermées positivement invariantes.

Dans le cas où $T$ est un homéomorphisme, on a

$$\begin{aligned} T^{-1}(\omega(x)) &= \bigcap_{m\geq 0} T^{-1}(\overline{O^+(T^m(x))}) \\ &= \bigcap_{m\geq 0} \overline{O^+(T^{m-1}(x))} \\ &= \bigcap_{m\geq -1} \overline{O^+(T^m(x))} \subset \omega(x). \end{aligned}$$

□

**Remarque.** Si $X$ est compact, alors $\omega(x) \neq \emptyset$. En effet, la suite $(\overline{O^+(T^m(x))})_{m\geq 0}$ est décroissante et formée de parties fermées non vides de $X$. Puisque $X$ est compact, l'intersection est non vide.

**Définition.** Si $T : X \to X$ est un homéomorphisme, on peut définir l'*ensemble $\alpha$-limite* $\alpha(x)$ d'un point $x \in X$ comme ensemble des points d'accumulation de la suite $(T^{-n}(x))_{n\geq 0}$. C'est une partie fermée et invariante.

**Remarque.** Si $Y$ est une partie fermée et positivement invariante, alors pour tout $y \in Y$, on a $\omega(y) \subset Y$. En particulier, pour tout $x \in X$ et tout $y \in \omega(x)$, on a $\omega(y) \subset \omega(x)$. Dans le cas où $T$ est un homéomorphisme, on a $\alpha(y) \subset \omega(x)$ puisque $\omega(x)$ est invariant.

**Définition.** Un point $x \in X$ est *positivement récurrent* si $x \in \omega(x)$. Ceci signifie qu'il existe une suite $(n_i)_{\geq 0}$ d'entiers positifs telle que $\lim_{i\to+\infty} n_i = +\infty$ et $\lim_{i\to+\infty} T^{n_i}(x) = x$.

**Remarques.**

1. On dit souvent récurrent au lieu de positivement récurrent.

2. Tout point périodique est récurrent.

PROPOSITION 1.1.3 : *L'ensemble* $\mathrm{Rec}^+(T)$ *des points positivement récurrents est positivement invariant (et invariant dans le cas où $T$ est inversible)*

*Preuve.* Supposons qu'il existe une suite $(n_i)_{\geq 0}$ d'entiers positifs telle que $\lim_{i\to+\infty} n_i = +\infty$ et $\lim_{i\to+\infty} T^{n_i}(x) = x$. Ceci implique que

$$\lim_{i\to+\infty} T^{n_i}(T(x)) = \lim_{i\to+\infty} T(T^{n_i}(x)) = T(x).$$

Dans le cas où $T$ est un homéomorphisme, on a

$$\lim_{i\to+\infty} T^{n_i}(T^{-1}(x)) = \lim_{i\to+\infty} T^{-1}(T^{n_i}(x)) = T^{-1}(x).$$

□

**Remarques.**

1. On verra dans le prochain chapitre que $\mathrm{Rec}^+(T)$ n'est pas nécessairement fermé.

2. Si $T$ est un homéomorphisme, on peut définir de façon similaire l'ensemble $\mathrm{Rec}^-(T)$ des points *négativement récurrents*, ce sont les points $x$ tels que $x \in \alpha(x)$. C'est également un ensemble fermé invariant. On verra également que l'on peut avoir $\mathrm{Rec}^+(T) \neq \mathrm{Rec}^-(T)$.

Concluons cette section par une notion encore plus faible de récurrence.

**Définition.** Un point $x \in X$ est *errant* s'il existe un voisinage $U$ de $x$ tel que $T^{-n}(U) \cap U = \emptyset$, pour tout $n \geq 1$. Dans le cas contraire, on dira que $x$ est *non errant.*

**Remarques.**

1. Bien évidemment les propriétés $T^{-n}(U) \cap U \neq \emptyset$ et $U \cap T^n(U) \neq \emptyset$ sont équivalentes. Cependant les dynamiciens préfèrent généralement la première formulation : pour toute partie $Y \subset X$ et tout entier $n \geq 0$, l'ensemble $T^{-n}(Y)$ est formé des points $x \in X$ qui sont dans $Y$ au temps $n$ sous l'action de la dynamique. De plus, si $Y$ est ouvert, il en est de même de $T^{-n}(Y)$, ce qui n'est pas nécessairement le cas de $T^n(Y)$). Un point $x$ est non errant si, pour tout voisinage $U$ de $x$, il existe un point $y \in U$ qui revient dans $U$ au bout d'un certain temps.

2. Dans le cas d'un homéomorphisme, l'errance "positive" ou "négative" coïncident, il n'y a pas d'ambiguité dans le terme "errance".

3. Remarquons que tout point positivement récurrent (et également négativement récurrent dans le cas d'un homéomorphisme) est non errant.

4. Une partie ouverte $U$ telle que $T^{-n}(U) \cap U = \emptyset$, pour tout $n \geq 1$, est généralement appelée *domaine errant.*

PROPOSITION 1.1.4 : *L'ensemble $\Omega(T)$ des points non errants est fermé et positivement invariant (et invariant dans le cas où $T$ est un homéomorphisme).*

*Preuve.* Si $x \notin \Omega(T)$, il existe un voisinage $U$ de $x$ tel que $T^{-n}(U) \cap U = \emptyset$, pour tout $n \geq 1$. Ceci implique que $U \cap \Omega(T) = \emptyset$. Nous venons de prouver que $X \setminus \Omega(T)$ est ouvert.

Soit $x \in \Omega(T)$ et $U$ un voisinage de $T(x)$. L'ensemble $T^{-1}(U)$ étant un voisinage de $x$, il existe $n \geq 1$ tel que

$$T^{-n}\left(T^{-1}(U)\right) \cap T^{-1}(U) = T^{-n-1}(U) \cap T^{-1}(U) \neq \emptyset.$$

Ceci implique que

$$\emptyset \neq T\left(T^{-n-1}(U) \cap T^{-1}(U)\right) \subset T^{-n}(U) \cap U.$$

Ainsi, on a $T(x) \in \Omega(T)$.

La remarque **ii)** précédant la proposition 1.1.4 nous dit que $\Omega(T)$ est également négativement invariant si $T$ est un homéomorphisme. □

Pour résumer cette section, on a les inclusions suivantes, en notant $\mathrm{Fix}(T)$ l'ensemble des points fixes de $T$ et $\mathrm{Per}(T)$ l'ensemble des points périodiques :

$$\mathrm{Fix}(T) \subset \mathrm{Per}(T) \subset \mathrm{Rec}^{\pm}(T) \subset \Omega(T).$$

Dans le cas d'un espace métrique, il existe une notion encore plus faible de récurrence que la non errance, il s'agite de la *récurrence par chaîne* (voir exercice **).

## 1.2 Conjugaison, facteur.

Commençons par introduire une notion d'isomorphisme entres systèmes dynamiques.

**Définition.** On dira que deux applications continues $T : X \to X$ et $S : Y \to Y$, définies sur des espaces topologiques $X$ et $Y$ sont *conjuguées* s'il existe un homéomorphisme $H : X \to Y$ tel que $H \circ T = S \circ H$. L'homéomorphisme $H$ est une *conjugaison* entre $T$ et $S$.

**Exemple.** Posons $S^1 = \{z \in \mathbf{C} \mid |z| = 1\}$. L'application

$$\begin{aligned} \mathrm{Exp} : \mathbf{R} &\to S^1 \\ x &\mapsto e^{2i\pi x} \end{aligned}$$

passe au quotient et défini un homéomorphisme naturel $H$ entre le groupe topologique $\mathbf{T}^1 = \mathbf{R}/\mathbf{Z}$ et $S^1$ tel que pour tout $x \in \mathbf{R}$ on ait $H(x + \mathbf{Z}) = e^{2i\pi x}$. Les applications

$$\begin{aligned} T : \mathbf{T}^1 &\to \mathbf{T}^1 \\ \widehat{x} &\mapsto 2\widehat{x} \end{aligned}$$

et

$$\begin{aligned} S : S^1 &\to S^1 \\ z &\mapsto z^2 \end{aligned}$$

sont conjuguées par $H$, les deux systèmes dynamiques sont isomorphes.

Proposition 1.2.1 : *Si $T : X \to X$ et $S : Y \to Y$ sont conjuguées par $H : X \to Y$, alors pour tout $n \geq 0$ on a $H \circ T^n = S^n \circ H$. Si l'une des applications $T$ ou $S$ est un homéomorphisme, il en est de même de l'autre et on a $H \circ T^k = S^k \circ H$ pour tout $k \in \mathbf{Z}$.*

*Preuve.* Le fait que $H \circ T^n = S^n \circ H$ pour tout $n \geq 0$ se prouve aisément par récurrence. La relation est évidente pour $n = 0$ ; si elle est vraie au rang $n - 1$, elle l'est également au rang $n$ car :

$$H \circ T^n = H \circ T^{n-1} \circ T = S^{n-1} \circ H \circ T = S^{n-1} \circ S \circ H = S^n \circ H.$$

On peut écrire $S = H \circ T \circ H^{-1}$. Si $T$ est un homéomorphisme, c'est également le cas de $S$ et on a

$$S^{-1} = (H^{-1})^{-1} \circ T^{-1} \circ H^{-1} = H \circ T^{-1} \circ H^{-1}.$$

L'homéomorphisme $H$ conjuguant $T^{-1}$ à $S^{-1}$, conjugue également $T^{-n}$ à $S^{-n}$ pour tout $n \geq 0$. $\square$

Le point important est que $H$ envoie les orbites de $T$ sur les orbites de $S$. On en déduit :

Proposition 1.2.2 : *Supposons que $H : X \to Y$ conjugue $T : X \to X$ à $S : Y \to Y$.*

**i)** *Pour tout $x \in X$, on a $H(\omega_T(x)) = \omega_S(H(x))$.*

**ii)** *L'image par $H$ des ensembles $\mathrm{Per}(T)$, $\mathrm{Rec}^+(T)$, $\Omega(T)$ sont égaux respectivement à $\mathrm{Per}(S)$, $\mathrm{Rec}^+(S)$, $\Omega(S)$. De plus, pour tout $x \in \mathrm{Per}(T)$, les points $x$ et $H(x)$ ont la même période.*

**iii)** *Si $T$ est un homéomorphisme, alors $H(\mathrm{Rec}^-(T)) = \mathrm{Rec}^-(S)$. De plus, $H(\alpha_T(x)) = \alpha_S(T(x))$ pour tout $x \in X$.*

*Preuve.* Commençons par prouver **i)**. Si $x' \in \omega_T(x)$, alors il existe une suite $(n_i)_{\geq 0}$ d'entiers positifs telle que $\lim_{i\to+\infty} n_i = +\infty$ et $\lim_{i\to+\infty} T^{n_i}(x) = x'$. On en déduit que

$$H(x') = \lim_{i\to+\infty} H\left(T^{n_i}(x)\right) = \lim_{i\to+\infty} (S^{n_i}(H(x))$$

ce qui signifie que $H(x') \in \omega_S(H(x))$. Réciproquement, si $y' \in \omega_S(y)$, alors $H^{-1}(y') \in \omega_T(H^{-1}(y))$.

On déduit de **i)** que $H$ envoie $\mathrm{Rec}^+(T)$ sur $\mathrm{Rec}^+(S)$. Le fait que l'image par $H$ d'un point périodique de $T$ est un point périodique de $S$ de même période provient de :

$$T^n(x) = x \Leftrightarrow H(T^n(x)) = H(x) \Leftrightarrow S^n(H(x)) = H(x).$$

Le lecteur vérifiera que l'image par $H$ d'un domaine errant de $T$ est un domaine errant de $S$ et par conséquence que $H$ envoie $\Omega(T)$ sur $\Omega(S)$. Il établira l'assertion **iii)**. $\square$

**Définition.** Soient $T : X \to X$ et $S : Y \to Y$ deux applications continues. L'application $S$ est un *facteur* de $T$ s'il existe une application surjective $H : X \to Y$ telle que $H \circ T = S \circ H$. L'application $H$ est une *semi-conjugaison*.

**Exemples.**

1. Fixons $\widehat{a} \in \mathbf{T}^1$. La rotation

$$\begin{aligned} S\ :\ & \mathbf{T}^1 \to \mathbf{T}^1 \\ & \widehat{x} \mapsto \widehat{x} + \widehat{a} \end{aligned}$$

est un facteur de

$$\begin{aligned} T\ :\ & \mathbf{T}^2 \to \mathbf{T}^2 \\ & (\widehat{x}_1, \widehat{x}_2) \mapsto (\widehat{x}_1 + \widehat{a}, \widehat{x}_1 + \widehat{x}_2) \end{aligned}.$$

La semi-conjugaison est la projection

$$\begin{aligned} H\ :\ & \mathbf{T}^2 \to \mathbf{T}^1 \\ & (\widehat{x}_1, \widehat{x}_2) \mapsto \widehat{x}_1 \end{aligned}.$$

2. En considérant un facteur $S$ de $T$, on perd de l'information sur la dynamique $T$. L'application

$$\begin{aligned} S\ :\ & \{0\} \to \{0\} \\ & 0 \mapsto 0 \end{aligned}$$

est un facteur universel, il ne donne aucune information. Le lecteur établira facilement le résultat suivant :

PROPOSITION 1.2.2: *Supposons que $H : X \to Y$ est une semi-conjugaison entre $T : X \to X$ et $S : Y \to Y$.*

**i)** *Pour tout $n \geq 0$ on a $H \circ T^n = S^n \circ H$.*

**ii)** *Pour tout $x \in X$, on a $H(\omega_T(x)) \subset \omega_S(H(x))$.*

**iii)** *L'image par $H$ des ensembles $\mathrm{Per}(T)$, $\mathrm{Rec}^+(T)$, $\Omega(T)$ sont inclus respectivement dans $\mathrm{Per}(S)$, $\mathrm{Rec}^+(S)$, $\Omega(S)$. De plus, pour tout $x \in \mathrm{Per}(T)$, la période de $H(x)$ divise celle de $x$.*

**iv)** *Si $T$ et $S$ sont des homéomorphismes, $H \circ T^k = S^k \circ H$, pour tout $k \in \mathbf{Z}$. Par conséquent, $H(\mathrm{Rec}^-(T)) \subset \mathrm{Rec}^-(S)$ et $H(\alpha_T(x)) \subset \alpha_S(H(x))$ pour tout $x \in X$.*

## 1.3 Propriétés des systèmes dynamiques.

Nous ne nous intéresserons ici, non pas aux propriétés des orbites d'un système dynamique donné, mais aux propriétés du système lui-même.

**Définition.** Nous dirons qu'une transformation continue $T : X \to X$ est *positivement transitive* si, pour toutes parties ouvertes non vides $U$ et $V$, il existe $n \geq 0$ tel que $U \cap T^{-n}(V) \neq \emptyset$. Si $T$ est un homéomorphisme, nous dirons que $T$ est *transitif* si, pour toutes parties ouvertes non vides $U$ et $V$, il existe $k \in \mathbf{Z}$ tel que $U \cap T^{-k}(V) \neq \emptyset$.

**Remarques.**

1. Dans le cas d'un homéomorphisme, la transitivité positive est plus forte que la transitivité. Par exemple l'homéomorphisme

$$\begin{aligned} T : \mathbf{Z} &\to \mathbf{Z} \\ x &\mapsto x+1 \end{aligned}$$

est transitif mais pas positivement transitif. Néanmoins, dans le cas où $X$ n'a pas de point isolé, les deux notions coïncident (voir exercice **).

2. Une façon équivalente pour exprimer la transitivité positive de $T$ est de demander que $\bigcup_{n\geq 0} T^{-n}(V)$ soit dense, pour toute partie ouverte non vide $V$. De façon similaire, un homéomorphisme $T$ est transitif si $\bigcup_{k\in\mathbf{Z}} T^{-k}(V)$ est dense, pour toute partie ouverte non vide $V$.

3. Dans le cas où $T$ n'est pas inversible, on dit également transitive au lieu de transitivité positive, en l'absence s'ambiguité.

Nous allons donner une définition équivalente de la transitivité. Rappelons quelques définitions de topologie.

**Definition.** Un espace topologique est un *espace de Baire* si l'intersection de toute famille dénombrable de parties ouvertes denses est dense.

**Remarques.**

1. Un espace topologique compact, ou un espace métrique complet, est un espace de Baire.

2. Un ensemble $G_\delta$ d'un espace topologique est l'intersection d'une famille dénombrable de parties ouvertes. Une partie $X$ d'un espace de Baire est *résiduelle* si elle contient un ensemble $G_\delta$ dense. Remarquons que l'intersection d'une famille dénombrable de parties résiduelles est elle-même résiduelle.

**Définition.** Un espace topologique est *séparable* s'il existe une base dénombrable de parties ouvertes $(O_i)_{i\in I}$ : toute partie ouverte est une union d'ensembles $O_i$. C'est par exemple le cas d'un espace métrique compact ou de $\mathbf{R}^n$.

PROPOSITION 1.3.1 : *Soit $X$ un espace de Baire séparable et $T : X \to X$ une transformation continue. Les propriétés suivantes sont équivalentes :*

**i)** *$T$ est positivement transitif ;*

**ii)** *il existe $x$ tel que $\omega(x) = X$ ;*

**iii)** *l'ensemble des $x \in X$ tels que $\omega(x) = X$ est résiduel.*

*Preuve.* Commençons par prouver que **ii)** implique **i)**. Donnons nous deux parties ouvertes $U$ et $V$. Le fait que $\omega(x) = X$ implique qu'il existe $n \geq 0$ tel que $T^n(x) \in U$ et qu'il existe $n' \geq n$ tel que $T^{n'}(x) \in V$. Le point $T^n(x)$ appartient donc à $U \cap T^{n-n'}(V)$. Il existe donc $n'' \geq 0$ tel que $U \cap T^{-n''}(V) \neq \emptyset$. Remarquons que nous n'avons besoin d'aucune hypothèse sur $X$ pour prouver cette implication.

Puisqu'une partie résiduelle d'un espace de Baire est non vide, on sait que **iii)** implique **ii)** Il reste donc à prouver que **i)** implique **iii)**. Considérons une base dénombrable $(U_i)_{i\in I}$ de la topologie de $X$. L'application $T$ étant positivement transitive, on sait que pour tout $i \in I$ et tout $m \geq 0$, l'ensemble ouvert

$$\bigcup_{n\geq m} T^{-n}(U_i) = \bigcup_{n\geq 0} T^{-n}(T^{-m}(U_i))$$

est dense. Ainsi,

$$Y = \bigcap_{i\in I,\, m\geq 0} \left( \bigcup_{n\geq m} T^{-n}(U_i) \right)$$

est un ensemble $G_\delta$ dense. Montrons que c'est exactement l'ensemble des points $x$ tels que $\omega(x) = X$. Un point $x$ appartient à $Y$ si et seulement si pour tout $i \in I$ et tout $m \geq 0$ il existe $n \geq m$ tel que $T^n(x) \in U_i$. Chaque $U_i$ étant ouvert on doit avoir $x \in Y$ si $\omega(x) = X$. Mais la réciproque est également vraie car $(U_i)_{i\in I}$ est une base de la topologie. En effet, pour tout $x' \in X$ et tout voisinage $U$ of $x'$, on peut trouver $i \in I$ tel que $U_i \subset U$. Ainsi, pour tout $m \geq 0$, il existe $n \geq m$ tel que $T^n(x) \in U$. □

**Exemples** On verra dans la prochaine section que

$$\begin{aligned} T_1 \ : \mathbf{T}^1 &\to \mathbf{T}^1 \\ \widehat{x} &\mapsto 2\widehat{x} \end{aligned}$$

est positivement transitive et que

$$\begin{aligned} T_2 \ : \mathbf{T}^1 &\to \mathbf{T}^1 \\ \widehat{x} &\mapsto \widehat{x} + \widehat{a} \end{aligned}$$

est positivement transitive si et seulement si $\widehat{a} \notin \mathbf{Q}/\mathbf{Z}$.

On a un résultat analogue pour la transitivité :

Proposition 1.3.2 : *Soit $X$ un espace de Baire séparable et $T \ : X \to X$ un homéomorphisme. Les propriétés suivantes sont équivalentes :*

**i)** *$T$ est transitif ;*

**ii)** *il existe $x$ tel que $\overline{O(x)} = X$ ;*

**iii)** *l'ensemble des $x \in X$ tels que $\overline{O(x)} = X$ est résiduel.*

*Preuve.* Commençons là-encore par prouver que **ii)** implique **i)**. Donnons nous deux parties ouvertes $U$ et $V$. Le fait que $\overline{O(x)} = X$ implique qu'il existe $k \in \mathbf{Z}$ et $k' \in \mathbf{Z}$ tels que $T^k(x) \in U$ et $T^{k'}(x) \in V$. Le point $T^k(x)$ appartient à $U \cap T^{k-k'}(V)$.

Là-encore, on sait que **iii)** implique **ii)** et il reste donc à prouver que **i)** implique **iii)**. Considérons une base dénombrable $(U_i)_{i\in I}$ de la topologie de $X$. L'homéomorphisme $T$ étant transitif, on sait que pour tout $i \in I$, l'ensemble ouvert $\bigcup_{k\in\mathbf{Z}} T^{-k}(U_i)$ est dense. Ainsi,

$$Y = \bigcap_{i\in I} \bigcup_{k\in\mathbf{Z}} T^{-k}(U_i)$$

est un ensemble $G_\delta$ dense. Un point $x$ appartient à $Y$ si et seulement si son orbite $O(x)$ rencontre chaque $U_i$, c'est-à-dire si elle est dense dans $X$. $\square$

**Définition.** Une transformation continue $T : X \to X$ est *(positivement) topologiquement mélangeante* si, pour toutes parties ouvertes $U$ et $V$, il existe $m \geq 0$ tel que $U \cap T^{-n}(V) \neq \emptyset$ pour tout $n \geq m$.

**Exemples** Il s'agit d'une propriété plus forte que la transitivité. Nous verrons que l'application

$$\begin{aligned} T_1 : \mathbf{T}^1 &\to \mathbf{T}^1 \\ \widehat{x} &\mapsto 2\widehat{x} \end{aligned}$$

est topologiquement mélangeante mais que

$$\begin{aligned} T_2 : \mathbf{T}^1 &\to \mathbf{T}^1 \\ \widehat{x} &\mapsto \widehat{x} + \widehat{a} \end{aligned}$$

ne l'est pas, même si $\widehat{a} \notin \mathbf{Q}/\mathbf{Z}$.

Nous allons introduire une dernière notion qui est également plus forte que la transitivité.

PROPOSITION 1.3.3 : *Soit $T : X \to X$ une application continue sur un espace topologique. Les conditions suivantes sont équivalentes :*

**i)** *pour tout $x \in X$, on a $\overline{O^+(x)} = X$;*

**ii)** *pour tout $x \in X$, on a $\omega(x) = X$;*

**iii)** *si $Y$ est une partie fermée positivement invariante, alors $Y = \emptyset$ ou $Y = X$.*

*Si ces conditions sont vérifiées, on dira que $T$ est* positivement minimale.

*Preuve.* Montrons d'abord que **i)** ou **ii)** implique **iii)**. Supposons que $Y$ est une partie fermée positivement invariante. Si $Y \neq \emptyset$, on peut choisir $x \in Y$. Chacun des ensembles $\overline{O^+(x)}$ et $\omega(x)$ étant contenu dans $Y$, on en déduit que si **i)** ou **ii)** est vérifiée, alors on a $X \subset Y$.

Pour montrer que **iii)** implique **i)**, il suffit de remarquer que $\overline{O^+(x)}$ est une partie fermée non vide positivement invariante. Pour prouver que **i)** implique **ii)**, écrivons

$$\omega(x) = \bigcap_{m \geq 0} \overline{O^+(T^m(x))} = \bigcap_{m \geq 0} X = X.$$

$\square$

**Remarque.** Un système dynamique positivement minimal est positivement transitif.

**Exemples.** Nous verrons que

$$\begin{aligned} T_1 : \mathbf{T}^1 &\to \mathbf{T}^1 \\ \widehat{x} &\mapsto 2\widehat{x} \end{aligned}$$

n'est pas positivement minimale mais que

$$\begin{aligned} T_2 : \mathbf{T}^1 &\to \mathbf{T}^1 \\ \widehat{x} &\mapsto \widehat{x} + \widehat{a} \end{aligned}$$

est positivement minimale si $\widehat{a} \notin \mathbf{Q}/\mathbf{Z}$.

On a une définition analogue pour les homéomorphismes. La preuve du résultat qui suit est similaire à celle de la proposition 1.3.3.

Proposition 1.3.4 : *Soit $T : X \to X$ un homéomorphisme sur un espace topologique. Les propriétés suivantes sont équivalentes*

**i)** *pour tout $x \in X$, on a $\overline{O(x)} = X$;*

**ii)** *si $Y$ est une partie fermée globalement invariante, alors $Y = \emptyset$ ou $Y = X$.*

*Si ces conditions sont vérifiées, on dira que $T$ est* minimal.

**Remarques.**

1. Pour un homéomorphisme, la minimalité positive est plus forte que la minimalité. Par exemple,

$$\begin{aligned} T : \mathbf{Z} &\to \mathbf{Z} \\ x &\mapsto x+1 \end{aligned}$$

est minimal mais pas positivement minimal. Néanmoins, si $X$ est compact, les deux notions coïncident (voir exercice **).

2. Un résultat de Gottshalk nous dit qu'il n'existe pas d'homéomorphisme positivement minimal sur $\mathbf{R}^n$, ni même sur tout espace localement compact non compact. On montre facilement qu'il n'y a pas d'homéomorphisme minimal sur $\mathbf{R}$ mais on sait également, d'après un théorème de point fixe dû à Brouwer, qu'il n'y a pas d'homéomorphisme minimal sur $\mathbf{R}^2$. L'existence ou non d'homéomorphisme minimal sur $\mathbf{R}^n$, est par contre un problème ouvert si $n \geq 3$.

3. Là-encore, on parlera souvent de minimalité, au lieu de minimalité positive dans le cas d'un système dynamique non inversible, en l'absence d'ambiguité.

**Définition.** Soit $T : X \to X$ une application continue sur un espace topologique. Une partie fermée positivement invariante $Y \subset X$ est *positivement minimale* si la transformation restreinte $F|_Y : Y \to Y$ est positivement minimale.

La proposition qui suit est une conséquence immédiate de la proposition 1.3.3

Proposition 1.3.5 : *Les conditions suivantes sont équivalentes :*

**i)** *$Y$ est positivement minimal ;*

**ii)** *pour tout $x \in Y$, on a $\overline{O^+(x)} = Y$;*

**iii** *pour tout $x \in Y$, on a $\omega(x) = Y$;*

**iv)** *$Y$ est un élément minimal, pour l'inclusion, dans l'ensemble des parties fermées non vides et positivement invariantes de $X$.*

Similairement, si $T : X \to X$ est un homéomorphisme, une partie fermée invariante $Y \subset X$ est *minimale* si l'homéomorphisme restreint $F|_Y : Y \to Y$ est minimal et on a :

Proposition 1.3.6 : *Les conditions suivantes sont équivalentes :*

**i)** *$Y$ est minimal ;*

**ii)** *pour tout $x \in Y$, on a $\overline{O(x)} = Y$;*

**ii)** *$Y$ est un élément minimal, pour l'inclusion, dans l'ensemble des parties fermées non vides et invariantes de $X$.*

Énonçons maintenant le résultat important suivant :

Proposition 1.3.7 : *Soit $T : X \to X$ une transformation continue sur un espace topologique compact. Il existe alors une partie fermée positivement minimale $Y \subset X$.*

*Preuve.* Notons $\mathcal{F}$ l'ensemble des parties fermées non vides positivement invariantes. Il n'est pas vide car il contient $X$. Remarquons qu'il est inductif pour l'inclusion. En effet, si $(Y_i)_{i\in I}$ est une famille totalement ordonnée d'éléments de $\mathcal{F}$, l'ensemble $Y = \bigcap_{i\in I} Y_i$ est fermé, positivement invariant et vérifie $Y \subset Y_i$ pour tout $i \in I$. Pour montrer qu'il est dans $\mathcal{F}$, il reste à vérifier qu'il n'est pas vide. C'est une conséquence immédiate de la compacité de $X$ et du fait que pour toute partie finie $J \subset I$, on a $Y = \bigcap_{i\in J} Y_i \neq \emptyset$ (puisque la famille est totalement ordonnée). Il ne reste plus qu'à appliquer le lemme de Zorn pour obtenir un élement minimal de $\mathcal{F}$. □

**Remarques.**

1. Il n'y a pas de partie positivement minimale pour

$$\begin{aligned} T\ : \mathbf{Z} &\to \mathbf{Z} \\ x &\mapsto x+1 \end{aligned}.$$

L'ensemble $\mathcal{F}$ est totalement ordonné, il est constitué de $\mathbf{Z}$ et des ensembles $Y_i = \{i, i+1, \ldots\}$. Remarquons que $\bigcap_{i\in\mathbf{Z}} Y_i = \emptyset$.

2. Une preuve analogue à celle de la proposition 1.3.7 nous dit qu'il existe une partie fermée invariante minimale dans le cas où $T$ est un homéomorphisme d'un espace topolgique compact $X$.

Concluons ce chapitre par le résultat suivant, souvent appelé *Théorème de récurrence de Birkhoff* :

Proposition 1.3.8 : *Si $X$ est compact, toute application continue $T : X \to X$ a un point positivement récurrent.*

*Preuve.* Considérons une partie positivement minimale $Y \subset X$ puis choisissons $x \in Y$. □

## 1.4 Les exemples classiques

### Le décalage de Benouilli unilatéral

Soit $A$ un ensemble (alphabet) fini de cardinal $p \geq 2$. On note $X = A^{\mathbf{N}}$ l'ensemble des suites $x = (x_n)_{n\geq 0}$ dans $A$. Pour tous $x$ et $y$ dans $X$, on note $N(x, y)$ le premier entier $n \geq 0$ tel que $x_n \neq y_n$ (si $x = y$, on pose $N(x, y) = +\infty$). Fixons $\alpha \in ]0, 1[$ et définissons pour tous $x$, $y$ dans $X$ :

$$\begin{cases} d_\alpha(x, y) = \alpha^{N(x,y)} & \text{si } x \neq y, \\ d_\alpha(x, y) = 0 & \text{si } x = y. \end{cases}$$

On obtient ainsi une distance sur $X$. En fait, le seul fait non trivial à vérifier est l'inégalité triangulaire : on a

$$d_\alpha(x, z) \leq \max(d_\alpha(x, y), d_\alpha(y, z))$$

car

$$N(x, z) \geq \min(N(x, y), N(y, z)).$$

Pour tout mot $w = (w_0, \ldots, w_{m-1}) \in A^m$ et tout $n_0 \geq 0$ on peut définir le cylindre

$$C_w^{n_0} = \{x = (x_n)_{n \in \mathbf{N}} \in A^{\mathbf{N}} \mid x_{n+n_0} = w_n \text{ pour tout } n \in \{0, \ldots, m-1\}\}.$$

Proposition 1.4.1 : *Les parties ouvertes de $(X, d_\alpha)$ sont les unions de cylindres.*

*Preuve.* Tout cylindre $C_w^{n_0}$ est ouvert. En effet, posons $N = n_0 + m - 1$. On a

$$x \in C_w^{n_0} \text{ et } N(x, y) > N \Rightarrow y \in C_w^{n_0}.$$

La boule $B(x, \alpha^N)$ est donc incluse dans $C_w^{n_0}$, si $C_w^{n_0}$.

Réciproquement, pour tout $x \in X$ et tout $N \geq 0$, on a $B(x, \alpha^N) = C_w^0$, où $w = (x_0, \ldots, x_N)$. □

La proposition précédente nous dit que la topologie de $(X, d_\alpha)$ est la topologie produit, si $A$ est muni de la topologie discrète. Une suite $(x^m)_{m \geq 0}$ converge vers $x$ si, pour tout $N \geq 0$, il existe $M \geq 0$ tel que $x_n^m = x_n$, si $n \leq N$ et $m \geq M$. Puisque les cylindres sont en nombre dénombrable, $X$ est séparable. On vérifie sans peine que les cylindres sont également fermés. On peut en déduire que $X$ est totalement discontinu. On vérifie également qu'il n'y a pas de point isolé. L'espace $X$ est donc un *ensemble de Cantor* (tous les espaces compacts continus et sans point isolé sont homéomorphes et appelés ainsi).

Proposition 1.4.2 : *L'ensemble $X$ est compact.*

*Preuve.* C'est un cas particulier du théorème de Tychonov sur la compacité du produit d'espaces compacts. Il suffit ici d'appliquer le principe diagonal de Cantor. Soit $(x^m)_{m \geq 0}$ une suite d'éléments de $X$. Il existe $x_0 \in A$ et une sous-suite $(x^{\theta_0(m)})_{m \geq 0}$, tel que $x_0^{\theta_0(m)} = x_0$ pour tout $m \geq 0$. Il existe $x_1 \in A$ et une sous-suite $(x^{\theta_0 \circ \theta_1(m)})_{m \geq 0}$, telle que $x_1^{\theta_0 \circ \theta(m)} = x_1$ pour tout $m \geq 0$. On définit par récurrence une suite $x = (x_n)_{n \geq 0} \in X$ et une suite $(\theta_n)_{n \geq 0}$ de fonctions strictement croissantes $\theta_n : \mathbf{N} \to \mathbf{N}$ telles que $x_n^{\theta_0 \circ \ldots \circ \theta_n(m)} = x_n$ pour tout $m \geq 0$. On vérifie aisément que la fonction $\theta : m \mapsto \theta_0 \circ \ldots \circ \theta_m(m)$ est strictement croissante et que la sous-suite $(x^{\theta(m)})_{m \geq 0}$ converge vers $x$. □

Proposition 1.4.3 : *L'application*

$$\begin{aligned} \sigma &: X \to X \\ (x_n)_{n \geq 0} &\mapsto (x_{n+1})_{n \geq 0} \end{aligned}$$

*est continue, on l'appelle le* décalage de Bernouilli unilatéral.

*Preuve.* Remarquons que $N(x, y) \geq N \Rightarrow N(\sigma(x), \sigma(y)) \geq N - 1$. □

**Remarques.**

1. L'application $\sigma$ a exactement $p$ points fixes, qui sont les suites constantes. Plus généralement $\sigma^q$ a $p^q$ points fixes obtenus par concaténation d'un même mot de longueur $q$ dans l'alphabet $A$. Ce sont les points périodiques de période $q'|q$.

2. Toute suite $x = (x_n)_{n\geq 0} \in X$ est limite d'une suite $(x^m)_{m\geq 1}$ d'éléments de X qui sont tous périodiques. Il suffit de prendre pour $x^m$ la suite qui est périodique de période $m$ et qui vérifie $x_n^m = x_n$ si $n < m$. Le fait que l'ensemble des points périodiques de $\sigma$ est dense implique que $\sigma$ n'a pas de point errant. En effet, $\Omega(\sigma)$ est fermé et contient $\mathrm{Per}(\sigma)$.

PROPOSITION 1.4.4 : *Le décalage de Bernouilli est transitif, et même mieux, il est topologiquement mélangeant*

*Preuve.* II suffit de prouver que si $C_w^{n_0}$ et $C_{w'}^{n'_0}$ sont deux cylindres quelquonques, il existe un entier $N$ tel que pour tout $n \geq N$, on a $C_w^{n_0} \cap \sigma^{-n}\left(C_{w'}^{n'_0}\right) \neq \emptyset$. Or on a $\sigma^{-n}\left(C_{w'}^{n'_0}\right) = C_{w'}^{n'_0+n}$. Notons $w = (w_0, \ldots, w_n)$ et $w' = (w'_0, \ldots, w'_{n'})$. Si $n$ est assez grand, les ensembles $(n_0, \ldots, n_0 + m - 1)$ et $(n + n'_0, \ldots, n + n'_0 + m' - 1)$ sont disjoints et par conséquent on a $C_w^{n_0} \cap \sigma^{-n}\left(C_{w'}^{n'_0}\right) \neq \emptyset$. □

Pour finir, remarquons que tout point a $(\sharp A)^n$ antécédents par $\sigma^n$ et que pour tout $x \in X$, l'ensemble $\bigcup_{n\geq 0} \sigma^{-n}(\{x\})$ est dense. On vérifiera que cette propriété implique la transitivité.

### Le décalage de Bernouilli bilatéral

On garde le même alphabet $A$, mais cette fois ci $X = A^{\mathbf{Z}}$ est l'ensemble des suites bilatérales $x = (x_k)_{k\in\mathbf{Z}}$ dans $A$. Pour tous $x$ et $y$ dans $X$, on note $N(x, y)$ le premier entier $n \geq 0$ tel que $x_n \neq y_n$ ou $x_{-n} \neq y_{-n}$ (si $x = y$, on pose $N(x, y) = +\infty$). On fixe $\alpha \in ]0, 1[$ et on définit pour tous $x$, $y$ dans $X$ :

$$\begin{cases} d_\alpha(x, y) = \alpha^{N(x,y)} & \text{si } x \neq y, \\ d_\alpha(x, y) = 0 & \text{si } x = y. \end{cases}$$

On obtient là-encore une distance sur $X$. Comme dans le cas unilatéral, pour tout mot $w = (w_0, \ldots, w_{m-1}) \in A^m$ et tout $k_0 \in \mathbf{Z}$ on définit le cylindre

$$C_w^{k_0} = \{x = (x_k)_{k\in\mathbf{Z}} \in A^{\mathbf{Z}} \mid x_{n+k_0} = w_n \text{ pour tout } n \in \{0, \ldots, m-1\}\}.$$

On obtient ainsi une base dénombrable de la topologie de $d_\alpha$. On démontre là-encore que $X$ est compact.

PROPOSITION 1.4.5 : *L'application*

$$\begin{aligned} \sigma \ &: X \to X \\ (x_k)_{k\in\mathbf{Z}} &\mapsto (x_{k+1})_{k\in\mathbf{Z}} \end{aligned}$$

*est un homéomorphisme. On l'appelle le* décalage de Bernouilli bilatéral.

*Preuve.* Ici encore

$$N(x, y) \geq N \Rightarrow N(\sigma(x), \sigma(y)) \geq N - 1 \text{ et } N(\sigma^{-1}(x), \sigma^{-1}(y)) \geq N - 1.$$

□

L'homéomorphisme $\sigma^q$ a $p^q$ points fixes obtenus par concaténation du même mot de longueur $q$ dans l'alphabet $A$. Comme dans le cas unilatéral, on a $\overline{\mathrm{Per}(\sigma)} = \Omega(\sigma) = X$ et on démontre de façon analogue :

PROPOSITION 1.4.6 : *Le décalage de Bernouilli est positivement (et négativement) topologiquement mélangeant*

**Les rotations de $\mathbf{T}^1$**

On fixe $a \in \mathbf{R}$ et on pose $\widehat{a} = a + \mathbf{Z}$. On rappelle que la translation $T_a \ : x \mapsto x + a$ relève la rotation $T_{\widehat{a}} \ : x \mapsto \widehat{x} + \widehat{a}$.

PROPOSITION 1.4.7 : *Si $a = \dfrac{p}{q}$ est rationnel, écrit sous forme irréductible, alors tout point $\widehat{x} \in \mathbf{T}^1$ est un point périodique de $T_{\widehat{a}}$, de période $q$.*

*Preuve.* Soit $\widehat{x} \in \mathbf{T}^1$. On écrit $\widehat{x} = x + \mathbf{Z}$, où $x \in \mathbf{R}$, et on a

$$T_{\widehat{a}}^q(\widehat{x}) = T_a^q(x) + \mathbf{Z} = x + p + \mathbf{Z} = \widehat{x},$$

ce qui implique que $\widehat{x}$ est un point périodique de $T_{\widehat{a}}$, de période $q'|q$. On en déduit que $T_a^{q'}(x)+\mathbf{Z} = \widehat{x}$. Ceci signifie qu'il existe $p' \in \mathbf{Z}$ tel que $x + q'a = x + p'$. Ainsi, $a = \dfrac{p'}{q'}$, ce qui implique que $p' = p$ et $q' = q$. □

PROPOSITION 1.4.8 : *Si $a$ est irrationnel, alors $T_{\widehat{a}}$ est positivement et négativement minimal.*

Commençons par une démonstration classique.

LEMMA 1.4.9 : *L'ensemble $\mathbf{Z} + a\mathbf{Z}$ est dense dans $\mathbf{R}$.*

*Preuve.* L'ensemble $\mathbf{Z}+a\mathbf{Z}$ est un sous groupe de $\mathbf{R}$. Commençons par prouver, par l'absurde, que

$$\varepsilon_0 = \inf\left((\mathbf{Z} + a\mathbf{Z})\cap]0, +\infty[\right) = 0.$$

Si $\varepsilon_0 > 0$, il existe au plus un point de $\mathbf{Z} + a\mathbf{Z}$ dans $[\varepsilon_0, 2\varepsilon_0[$. En effet, la différence entre deux tels nombres nous donnerait un point appartenant à la fois à $\mathbf{Z} + a\mathbf{Z}$ et à $]0, \varepsilon_0[$. On en déduit que $\varepsilon_0 \in \mathbf{Z} + a\mathbf{Z}$. On peut écrire

$$1 = q\varepsilon_0 + r, \ \ a = p\varepsilon_0 + r'$$

où $r$ et $r'$ appartiennent à $[0, \varepsilon_0[$ et où $p$ et $q$ sont entiers. Puisque $r$ et $r'$ appartiennent à $\mathbf{Z}+a\mathbf{Z}$, ils sont nuls. Ceci implique que $a = \dfrac{p}{q}$. On a donc une contradiction.

Pour conclure le lemme, on fixe $x \in \mathbf{R}$ et $\varepsilon > 0$, on choisit $y \in (\mathbf{Z} + a\mathbf{Z})\cap]0, \varepsilon[$ puis l'entier $k \in \mathbf{Z}$ tel que $ky \le x < (k + 1)y$. □

LEMMA 1.4.10 : *L'ensemble $\mathbf{Z} + a\mathbf{N}$ est dense dans $\mathbf{R}$.*

*Preuve.* Le lemme 3.4.9 nous dit qu'il existe deux suites $(p_n)_{n\ge0}$ et $(q_n)_{n\ge0}$ dans $\mathbf{Z}$ telles que la suite $(p_n + q_n a)_{n\ge0}$ est décroissante et converge vers 0. Les deux suites $(p_n)_{n\ge0}$ et $(q_n)_{n\ge0}$ ne sont évidemment pas bornées.

Si $(q_n)_{n\geq 0}$ n'est pas majorée, on obtient immédiatement que $\mathbf{Z}+a\mathbf{N}$ rencontre chaque intervalle $]0,\varepsilon[$. On en déduit ensuite que $\mathbf{Z}+a\mathbf{N}$ est dense dans $[0,+\infty[$ en utilisant la division euclidienne, puis bien sûr dans chaque intervalle $[k,+\infty)$, $k\leq 0$.

Si $(q_n)_{n\geq 0}$ n'est pas minorée, alors pour tout $\varepsilon>0$, on peux trouver $n$ et $n'$ tels que

$$0<p_{n'}+q_{n'}a<p_n+q_na<\varepsilon \text{ et } q_{n'}<q_n.$$

On en déduit que

$$0<p_n-p_{n'}+(q_n-q_{n'})a<\varepsilon,$$

ce qui permet de conclure comme précédemment. $\square$

*Preuve de la proposition 1.4.8.* Fixons $\widehat{x}\in\mathbf{T}^1$ et écrivons $\widehat{x}=x+\mathbf{Z}$, où $x\in\mathbf{R}$. L'orbite $O^+(\widehat{x})$ n'est rien d'autre que la projection dans $\mathbf{T}^1$ de l'ensemble $x+\mathbf{Z}+a\mathbf{N}$ qui, on vient de le voir, est dense dans $\mathbf{R}$. Ainsi, $O^+(\widehat{x})$ est dense dans $\mathbf{T}^1$. $\square$

On peut donner également la preuve rapide suivante.

*Autre preuve de la proposition 1.4.8.* Puisque $\mathbf{T}^1$ est compact, il suffit de prouver que $T_{\widehat{a}}$ est minimal. Supposons qu'il existe un ensemble fermé invariant $X$ qui ne soit ni vide, ni égal à $\mathbf{T}^1$. L'homéomorphisme $T_{\widehat{a}}$ induit une bijection naturelle sur l'ensemble des composantes connexes de $\mathbf{T}^1\setminus X$. Si $\widehat{d}$ est la distance sur $\mathbf{T}^1$ induite par la norme usuelle de $\mathbf{R}$, nous savons que $T_{\widehat{a}}$ est une isométrie : pour tous $\widehat{x}$ et $\widehat{x}'$ dans $\mathbf{T}^1$, nous avons $\widehat{d}(T_{\widehat{a}}(\widehat{x}),T_{\widehat{a}}(\widehat{x}'))=\widehat{d}(\widehat{x},\widehat{x}')$. Fixons $\delta>0$ assez petit. Il y a alors un nombre fini (non nul) de composantes connexes de $\mathbf{T}^1\setminus X$ et elles sont permutées par $T_{\widehat{a}}$. Elles sont donc périodiques, ce qui implique que leurs extrémités sont des points périodiques de $T_{\widehat{a}}$. Il reste à prouver que $T_{\widehat{a}}$ n'a pas de point périodique. Raisonnons par l'absurde. Si $T^q_{\widehat{a}}(\widehat{x})=\widehat{x}$ et si $x\in\pi^{-1}(\{\widehat{x}\})$, alors il existe $p\in\mathbf{Z}$ tel que $x+qa=x+p$, ce qui implique que $a=p/q$, contredisant l'irrationalité de $a$. $\square$

Proposition 1.4.11 : *La rotation $T_{\widehat{a}}$ n'est pas topologiquement mélangeante.*

*Preuve.* Utilisons le fait que $T_{\widehat{a}}$ est une isométrie pour $\widehat{d}$. Supposons que $\widehat{x}$ et $\widehat{x}'$ soient distincts et posons $\varepsilon=\widehat{d}(\widehat{x},\widehat{x}')>0$. Considérons les boules $U=B(\widehat{x},\dfrac{\varepsilon}{4})$ et $U'=B(\widehat{x}',\dfrac{\varepsilon}{4})$. Que ce soit dans le cas rationnel ou irrationnel, on peut trouver $n$ arbitrairement grand tel que $d(T^n_{\widehat{a}}\widehat{x},\widehat{x})<\dfrac{\varepsilon}{4}$. Ceci implique que $U\cap T^{-n}(U')=\emptyset$. $\square$

La rotation d'angle $2\pi a$ :

$$\begin{aligned} R_a\ :\ & S^1\to S^1\\ & z\mapsto e^{2i\pi a}z\end{aligned}$$

est conjuguée à $T_{\widehat{a}}$ par l'homéomorphisme $H\ :\ \mathbf{T}^1\to S^1$ tel que $H(x+\mathbf{Z})=e^{2i\pi x}$. Toute propriété dynamique de $T_{\widehat{a}}$ est transportée par $H$ en une propriété de $R_a$.

### Les rotations de $\mathbf{T}^r$

Fixons une norme $\|\ \|$ sur $\mathbf{R}^r$ pour définir une distance $\widehat{d}$ sur $\mathbf{T}^r$. Fixons ensuite $a=(a_1,\dots,a_r)\in\mathbf{R}^r$ et considérons la translation $T_a\ :\ x\mapsto x+a$ qui relève la rotation $T_{\widehat{a}}\ :\ x\mapsto\widehat{x}+\widehat{a}$, où $\widehat{a}=a+\mathbf{Z}^r$. Ici, $T_{\widehat{a}}$ est également une isométrie.

Proposition 1.4.12 : *Tout point de $\mathbf{T}^r$ est un point positivement récurrent de $T_{\widehat{a}}$.*

*Preuve.* Le théorème de récurrence de Birkhoff nous dit qu'il existe au moins un point positivement récurrent $\widehat{x}_0$. Toute rotation $T_{\widehat{x}}$ commute avec $T_{\widehat{a}}$, c'est-à-dire conjugue $T_{\widehat{a}}$ à elle-même. Ainsi $T_{\widehat{x}}(\widehat{x}_0)$ est également un point positivement récurrent de $T_{\widehat{a}}$. Ceci implique bien évidemment la proposition. □

On a la généralisation suivante du cas uni-dimensionnel :

Proposition 1.4.13 : *Les assertions suivantes sont équivalentes :*

**i)** *$T_{\widehat{a}}$ est positivement transitive ;*

**ii)** *$T_{\widehat{a}}$ est positivement minimale ;*

**iii)** *les réels $1$, $a_1,\ldots$, $a_r$ sont rationnellement indépendants (il n'existe pas de $n+1$-uplet $(k_0,\ldots,k_r)\in\mathbf{Z}^{r+1}\setminus\{0\}$ tel que $k_0+k_1a_1+\ldots+k_ra_r=0$).*

*Preuve.* Les conditions **i)** et **ii)** sont équivalentes. En effet, l'argument de la preuve de la proposition 3.4.12 nous dit que si $\widehat{x}_0$ a un ensemble $\omega$-limite égal à $\mathbf{T}^r$, alors pour tout $\widehat{x}\in\mathbf{T}^r$, le point $T_{\widehat{x}}(\widehat{x}_0)$ a également un ensemble $\omega$-limite égal à $\mathbf{T}^r$ et par conséquent que la transitivité implique la minimalité.

Montrons maintenant par contraposée que **i)** implique **iii)**. Supposons qu'il existe $(k_0,\ldots,k_r)\in\mathbf{Z}^{r+1}\setminus\{0\}$ tel que $k_0+k_1a_1+\ldots+k_ra_r=0$. Considérons la forme linéaire

$$L\ :(x_1,\ldots,x_r)\mapsto\sum_{i=1}^r k_ix_i.$$

L'application, de $\mathbf{R}^r$ dans $\mathbf{T}^1$, qui à $x\in\mathbf{R}^r$ associe $L(x)+\mathbf{Z}$, est $\mathbf{Z}^r$-invariante : il existe donc une application continue $\widehat{L}\ :\mathbf{T}^r\to\mathbf{T}^1$ telle que, pour tout $x\in\mathbf{R}^r$, on ait $L(x)+\mathbf{Z}=\widehat{L}(x+\mathbf{Z}^r)$. Puisque $L$ est surjective, il en est de même de $\widehat{L}$. Remarquons maintenant que pour tout $x\in\mathbf{R}^r$, on a $L(x+a)=L(x)-k_0$, ce qui implique $\widehat{L}\circ T_{\widehat{a}}=\widehat{L}$. Ainsi, $\widehat{L}$ prend une valeur constante sur l'orbite $O(\widehat{x})$ d'un point $\widehat{x}$, ainsi que sur son adhérence. Ceci implique que $\omega(\widehat{x})\neq\mathbf{T}^r$. La rotation $T_{\widehat{a}}$ n'est pas positivement transitive. Nous monterons plus tard que **iii)** implique **i)** avec des arguments de théorie ergodique. □

En fait, on a :

Proposition 1.4.14 : *Si les nombres $1$, $a_1,\ldots$, $a_r$ sont rationnellement dépendants et si $s+1$ est la dimension du $\mathbf{Q}$-espace vectoriel engendré par $1$, $a_1,\ldots$, $a_r$, alors il existe un entier $q\geq 1$ tel que tout point $\widehat{x}$ appartient à un ensemble positivement et négativement minimal $X_{\widehat{x}}$ qui est l'union de $q$ tores de dimension $s$.*

### Les endomorphismes linéaires de $\mathbf{T}^1$

On considère dans ce paragraphe un endomorphisme $F\ :\widehat{x}\mapsto p\widehat{x}$ de $\mathbf{T}^1$, où $p\in\mathbf{Z}$. Si $p=0$ tout point est envoyé sur $0+\mathbf{Z}$ et la dynamique est triviale. La dynamique est également triviale si $p=1$, puisque $F=\mathrm{Id}_{\mathbf{T}^1}$ et guère plus si $p=-1$ puisque tout point est alors périodique de période 2 sauf $0+\mathbf{Z}$ et $1/2+\mathbf{Z}$ qui sont fixes. Remarquons que $F^2\ :\widehat{x}\mapsto p^2\widehat{x}$. On supposera donc que $p\geq 2$ dans cette section.

PROPOSITION 1.4.15 : *Tout point a $p$ antécédents par $F$ et $p^n$ par $F^n$. De plus l'orbite d'un point $\widehat{x} = x + \mathbf{Z}$ aboutit au point fixe $0 + \mathbf{Z}$ si et seulement s'il existe $k \in \mathbf{Z}$ et $n \geq 0$ tels que $x = \frac{k}{p^n}$.*

*Preuve.* L'équation $p\widehat{x} = \widehat{y}$, où $\widehat{y} = y + \mathbf{Z}$, a $p$ solutions : les projections dans $\mathbf{T}^1$ des réels $\frac{y}{p}$, $\frac{y}{p} + \frac{1}{p}$,..., $\frac{y}{p} + \frac{p-1}{p}$. Le reste de la proposition s'en déduit facilement. □

PROPOSITION 1.4.16 : *L'application $F$ a $p-1$ points fixes et plus généralement $F^q$ a $p^q - 1$ points fixes, si $q \geq 1$. De plus,* Per(F) *est dense dans $\mathbf{T}^1$.*

*Preuve.* Pour résoudre $p\widehat{x} = \widehat{x}$, on doit résoudre $(p-1)\widehat{x} = \widehat{0}$. On a vu que cette équation a $p-1$ solutions : les projections dans $\mathbf{T}^1$ des réels $0$, $\frac{1}{p-1}$,..., $\frac{p-2}{p-1}$. La fin de la première assertion s'en déduit immédiatement. Pour montrer que Per(F) est dense dans $\mathbf{T}^1$, il suffit de remarquer que l'ensemble des points de la forme $x = \frac{k}{p^q-1}$, $q \geq 0$, $k \in \mathbf{Z}$, est dense dans $\mathbf{R}$. □

On en déduit là-encore que tout point est non errant puisque les points périodiques sont denses. En fait on a :

PROPOSITION 1.4.17 : *L'application $F$ est topologiquement mélangeante..*

*Preuve.* On a une situation comparable à celle du décalage de Bernouilli unilatéral. Pour tout point $\widehat{y} \in \mathbf{T}^1$ l'ensemble $\bigcup_{n\geq 0} F^{-n}(\{\widehat{y}\})$ est dense $\mathbf{T}^1$. On a même mieux : pour toute partie ouverte $V$, il existe $N$ tel que $F^{-n}(\{\widehat{y}\}) \cap V \neq \emptyset$, si $n \geq N$. En effet, $F^{-n}(\{\widehat{y}\})$ est formé des projections des $p^n$ points $\frac{y}{p^n}$, $\frac{y}{p^n} + \frac{1}{p^n}$,..., $\frac{y}{p^n} + \frac{p^n-1}{p^n}$, où $\widehat{y} = y + \mathbf{Z}$. Ainsi tout point de $\mathbf{T}^1$ est à une distance au plus $\frac{1}{2p^n}$ d'un point de $F^{-n}(\{y\})$. Cette propriété implique bien évidemment que $F$ est topologiquement mélangeante. □

**Remarques.**

1. Ce qui précède nous dit que le système dynamique précédent est fortement dépendant des conditions initiales. Deux points peuvent être très proches et avoir des orbites qui divergent. C'est une situation radicalement différente de celle d'une rotation, où deux points proches restent proches au cours du temps sous l'action de la dynamique.

2. L'application

$$\begin{aligned} G \ : S^1 &\to S^1 \\ z &\mapsto z^p \end{aligned}$$

étant conjuguée à $F$, partage les mêmes propriétés dynamiques

PROPOSITION 1.4.18 *L'application $F$ est un facteur du décalage de Bernouilli*

$$\sigma \ : \{0, \ldots, p-1\}^{\mathbf{N}} \to \{0, \ldots, p-1\}^{\mathbf{N}}.$$

*Preuve.* On va démontrer que

$$\begin{aligned} H \ : \{0, \ldots, p-1\}^{\mathbf{N}} &\to \mathbf{T}^1 \\ x &\mapsto \sum_{n=0}^{\infty} \frac{x_n}{p^{n+1}} + \mathbf{Z} \end{aligned}$$

est une semi-conjugaison entre $F$ and $G$.

L'application

$$\begin{aligned} \widetilde{H} \ &: \{0,\dots,p-1\}^{\mathbf{N}} \to \mathbf{R} \\ x &\mapsto \sum_{n=0}^{\infty} \frac{x_n}{p^{n+1}} \end{aligned}$$

est bien définie et son image est l'intervalle $[0,1]$ (car tout point a un développement $p$-adique). Elle est et continue puisque

$$\begin{aligned} N(x,y) > N \Rightarrow |\widetilde{H}(x) - \widetilde{H}(y)| &\leq \sum_{n=N+1}^{\infty} \frac{p-1}{p^{n+1}} \\ &\leq (p-1)p^{-N-2}\frac{1}{1-p^{-1}} = p^{-N-1} \end{aligned}.$$

Remarquons maintenant que pour tout $x \in \{0,\dots,p-1\}^{\mathbf{N}}$, on a

$$p\widetilde{H}(x) = \widetilde{H}(\sigma(x)) + x_0.$$

Toutes ces propriétés impliquent que $H$ est continue, surjective et vérifie $H \circ \sigma = F \circ \sigma$. □

### Les endomorphismes linéaires de $\mathbf{T}^r$

Soit $A : \mathbf{R}^r \to \mathbf{R}^r$ une application linéaire dont la matrice dans la base canonique de $\mathbf{R}^r$ est à coefficients entiers. Puisque $A(\mathbf{Z}^r) \subset \mathbf{Z}^r$, on sait que $A(x) - A(y) \in \mathbf{Z}^r$ si $x - y \in \mathbf{Z}^r$. On en déduit que $A$ relève une application continue $\widehat{A} \ : \mathbf{T}^r \to \mathbf{T}^r$. Il s'agit d'un *endomorphisme linéaire*, c'est-à-dire d'un endomorphisme continu de $\mathbf{T}^r$. Il en est ainsi par exemple de

$$\begin{aligned} \widehat{A}_1 \ &: \mathbf{T}^1 \to \mathbf{T}^1 \\ \widehat{x} &\mapsto 2\widehat{x} \end{aligned}$$

et de

$$\begin{aligned} \widehat{A}_2 \ &: \mathbf{T}^2 \to \mathbf{T}^2 \\ (\widehat{x}_1, \widehat{x}_2) &\mapsto (2\widehat{x}_1 + \widehat{x}_2, \widehat{x}_1 + \widehat{x}_2) \end{aligned}.$$

Proposition 1.4.19 : *L'endomorphisme $\widehat{A}$ est surjectif si et seulement si* $\det(A) \neq 0$. *Dans ce cas, il préserve la mesure de Haar. De plus, tout point a exactement* $|\det(A)|$ *antécedents. L'endomorphisme $\widehat{A}$ est bijectif si et seulement si* $\det(A) = \pm 1$. *Dans ce cas $\widehat{A}^{-1}$ est un endomorphisme linéaire.*

*Preuve.* Dire que l'application $\widehat{A}$ est surjective signifie que pour tout $y \in \mathbf{R}^r$, il existe $x \in \mathbf{R}^r$ et $k \in \mathbf{Z}^r$ tels que $y = A(x) + k$. En d'autres termes, $\widehat{A}$ est surjective si et seulement si $\mathbf{R}^r = \bigcup_{k \in \mathbf{Z}^r} \mathrm{Im}(A) + k$.

Si $A$ est surjective, c'est-à-dire si $\det(A) \neq 0$, alors $\widehat{A}$ l'est aussi. Si $A$ n'est pas surjective, alors l'espace quotient $\mathbf{R}^r/\mathrm{Im}(A)$ est de dimension au moins 1 et n'est donc pas dénombrable, ce qui exclut que l'on puisse écrire $\mathbf{R}^r/\mathrm{Im}(A) = \bigcup_{k \in \mathbf{Z}^r} \{k + \mathrm{Im}(A)\}$.

Pour montrer que $\widehat{A}$ préserve la mesure de Haar $\mu$ dans le cas surjectif, il suffit de prouver que la mesure $\widehat{A}_*(\mu)$ est invariante par toute rotation. On doit donc prouver que pour tout point $\widehat{b} \in \mathbf{T}^r$ et tout ensemble borélien $\widehat{X}$, on a $\mu(\widehat{A}^{-1}(T_{\widehat{b}}^{-1}(\widehat{X}))) = \mu(\widehat{A}^{-1}(\widehat{X}))$. Puisque $\widehat{A}$ est un endomorphisme surjectif, on peut trouver $\widehat{a}$ tel que $\widehat{A}(\widehat{a}) = \widehat{b}$ et on a $\widehat{A}^{-1}(T_{\widehat{b}}^{-1}(\widehat{X})) = \widehat{T}_{\widehat{a}}^{-1}(A^{-1}(\widehat{X}))$, ce qui implique que $\mu(\widehat{A}^{-1}(T_{\widehat{b}}^{-1}(\widehat{X}))) = \mu(\widehat{A}^{-1}(\widehat{X}))$ puisque $\mu$ est invariante par $T_{\widehat{a}}$.

Observons d'abord que tous les points ont le même nombre d'antécédents puisque $\widehat{A}$ est un endomorphisme. Notons $p$ ce nombre et considérons le cube $Y = [-\delta, \delta]^r$ ainsi que le polyèdre $X = A^{-1}(Y)$. Si $\delta > 0$ est petit, $Y$ se projette injectivement sur un ensemble borélien $\widehat{Y}$ de mesure $(2\delta)^r$ et $X$ sur un ensemble borélien $\widehat{X}$ de mesure $|\det(A)|^{-1}(2\delta)^r$. L'ensemble $\widehat{A}^{-1}(\widehat{Y})$ est la réunion disjointe de $p$ translatés de $\widehat{X}$. Puisque $\widehat{A}$ préserve la mesure de Lebesgue, on en déduit que $p = |\det(A)|$.

On en déduit bien sûr que $\det(A) = \pm 1$ si $\widehat{A}$ est bijectif. Dans ce cas, la matrice de $A^{-1}$ dans la base canonique est à coefficients entiers et relève un endomorphisme linéaire de $\mathbf{T}^r$ qui n'est rien d'autre que $\widehat{A}^{-1}$. □

Proposition 1.4.20 : *Si l'une des valeurs propres de $A$ est racine de l'unité, alors il existe $n \geq 1$ tel que $\widehat{A}^n$ a une infinité de points fixes. Sinon, pour tout $n \geq 1$, l'ensemble* $\mathrm{Fix}(\widehat{A}^n)$ *est fini et a* $|\det(A^n - \mathrm{Id}_{\mathbf{R}^r})|$ *éléments.*

*Preuve.* Les points fixes de $\widehat{A}^n$ sont les antécédents de $0 + \mathbf{Z}$ par l'endomorphisme $\widehat{A}^n - \mathrm{Id}_{\mathbf{T}^r}$ et celui-ci est relevé par $A^n - \mathrm{Id}_{\mathbf{R}^r}$. □

**Remarque.** Si $\lambda_1, \ldots, \lambda_r$ sont les valeurs propres de $A$, alors

$$|\det(A^n - \mathrm{Id}_{\mathbf{R}^r})| = \prod_{i=1}^{r} |\lambda_i^n - 1|.$$

Ainsi, on a

$$\log(\sharp \mathrm{Fix}(\widehat{A}^n)) = \sum_{i=1}^{r} \log(|\lambda_i^n - 1|).$$

Dans le cas où aucune valeur propre n'est de module 1, on a

$$\lim_{n \to +\infty} \frac{1}{n} \log(\sharp \mathrm{Fix}(\widehat{A}^n)) = \sum_{i \in I} \log(|\lambda_i|),$$

où $i \in I$ si et seulement si $|\lambda_i| > 1$.

Proposition 1.4.21 : *Tout point de $\mathbf{Q}^r/\mathbf{R}^r$ est périodique ou prépériodique. Si les valeurs propres de $A$ ne sont pas racines de l'unité, alors tout point périodique appartient à $\mathbf{Q}^r/\mathbf{R}^r$.*

*Preuve.* Pour tout $q \geq 1$, l'ensemble $\frac{1}{q}\mathbf{Z}^r$ est positivement invariant par $A$. On en déduit que l'ensemble $\left(\frac{1}{q}\mathbf{Z}^r\right)/\mathbf{Z}^r$ est positivement invariant par $\widehat{A}$. Puisqu'il est fini, tout point $\widehat{x} \in \left(\frac{1}{q}\mathbf{Z}^r\right)/\mathbf{Z}^r$ est périodique ou prépériodique (il existe $k$ tel que $\widehat{A}(\widehat{x})$ est périodique). Remarquons maintenant que $\mathbf{Q}^r/\mathbf{Z}^r = \bigcup_{q \geq 1} \left(\frac{1}{q}\mathbf{Z}^r\right)/\mathbf{Z}^r$.

Supposons maintenant que $\widehat{x}$ est périodique et écrivons $\widehat{x} = x + \mathbf{Z}^r$, où $x \in \mathbf{R}^r$. Il existe $q \geq 1$ et $k \in \mathbf{Z}^r$ tels que $A^q(x) = x + k$. Si les valeurs propres de $A$ ne sont pas racines de l'unité, alors $A^q - \mathrm{Id}_r$ est bijective et les solutions de $A^q(x) - x = k$ appartiennent à $\mathbf{Q}^r$. □

## 2.1. Notations

Soit $F$ un homéomorphisme du cercle $\mathbf{T}^1$ et $f\ :\mathbf{R}\to\mathbf{R}$ un relèvement de $F$. L'application $x\mapsto f(x+1)-f(x)$ est à valeurs entières, et donc constante puisque $f$ est continue: il existe $k\in\mathbf{Z}$ tel que $f(x+1)=f(x)+k$, pour tout $x\in\mathbf{R}$. On sait d'autre part que $f|_{[0,1[}$ est injective et que son image ne contient pas deux points translatés d'un entier (puisque $F$ est injective). On en déduit que $k=\pm1$. Ainsi, deux cas sont possibles : on a $F_*=\mathrm{Id}_{\mathbf{R}}$ ou $F_*=-\mathrm{Id}_{\mathbf{R}}$. Dans le premier cas, $f$ est un homéomorphisme croissant et vérifie $f(x+1)=f(x)+1$, pour tout $x\in\mathbf{R}$ ; dans le second cas $f$ est un homéomorphisme décroissant et vérifie $f(x+1)=f(x)-1$, pour tout $x\in\mathbf{R}$. Le premier cas a lieu quand $F$ préserve l'orientation (son degré est 1), le second quand $F$ renverse l'orientation (son degré est $-1$). Nous nous intéresserons principalement au premier cas et noterons alors $\mathrm{Homeo}_+(\mathbf{T}^1)$ le groupe des homéomorphismes de $\mathbf{T}^1$ qui préservent l'orientation. Nous noterons $D^0(\mathbf{T}^1)$ le groupe de tous les relèvements des éléments de $\mathrm{Homeo}_+(\mathbf{T}^1)$, autrement dit le groupe des homéomorphismes croissants $f$ tels que $f-\mathrm{Id}_{\mathbf{R}}$ est périodique, de période 1 Ce groupe contient bien sûr les translations $T_a$, $a\in\mathbf{R}$. On écrira fréquemment $f+a$ au lieu de $T_a\circ f$, si $f\in D^0(\mathbf{T}^1)$. On munira $D^0(\mathbf{T}^1)$ de la distance suivante

$$d(f,g)=\max\left(\max_{x\in\mathbf{R}}|f(x)-g(x)|,\ \ \max_{x\in\mathbf{R}}|f^{-1}(x)-g^{-1}(x)|\right).$$

Proposition 2.1.1 : *On a les résultats suivants :*

**i)** *l'application $(f,g)\mapsto f\circ g$ est continue ;*

**ii)** *l'application $f\mapsto f^{-1}$ est continue ;*

**iii)** *l'espace $(D^0(\mathbf{T}^1),d)$ est complet.*

*Preuve.* L'assertion **ii)** est évidente puisque $d(f^{-1},g^{-1})=d(f,g)$. Prouvons **i)**. Supposons donc que $\lim_{n\to+\infty}f_n=f$ et $\lim_{n\to+\infty}g_n=g$ et montrons que $\lim_{n\to+\infty}f_n\circ g_n=f\circ g$. Fixons $\varepsilon>0$. Il existe $N\geq0$ tel que $d(f_n,f)\leq\varepsilon/2$ pour tout $n\geq N$. L'application $f$ est uniformément continue puisque c'est la somme de $\mathrm{Id}_{\mathbf{R}}$ et d'une application périodique. Il existe donc $\alpha>0$ tel que

$$|x-y|\leq\alpha\Rightarrow|f(x)-f(y)|\leq\varepsilon/2.$$

Il existe $N'\geq0$ tel que $d(g_n,g)\leq\alpha$ pour tout $n\geq N'$. On en déduit que pour tout $n\geq\max(N,N')$ on a

$$|f_n\circ g_n(x)-f\circ g(x)|\leq|f_n\circ g_n(x)-f\circ g_n(x)|+|f\circ g_n(x)-f\circ g(x)|\leq\varepsilon/2+\varepsilon/2=\varepsilon.$$

Le même argument nous dit que $|g_n^{-1}\circ f_n^{-1}(x)-g^{-1}\circ f^{-1}(x)|\leq\varepsilon$, si $n$ est assez grand. Pour prouver **iii)**, considérons une suite de Cauchy $(f_n)_{n\geq0}$ pour $d$. Chacune des suites $(f_n)_{n\geq0}$ et $(f_n^{-1})_{n\geq0}$ est une suite de Cauchy pour la distance $d'$, où

$$d'(f,g)=\max_{x\in\mathbf{R}}|f(x)-g(x)|,$$

et converge uniformément vers une application continue. Notons $f$ (resp. $g$) la limite de la première (resp. la seconde) Les égalités $f(x+1)=f(x)$ et $g(x+1)=g(x)+1$ sont également vraies. Il reste à prouver que $f$ et $g$ sont inversibles et que $g = f^{-1}$. Or la preuve de **i)** nous dit que

$$g \circ f = \lim_{n \to +\infty} f_n^{-1} \circ f_n = \mathrm{Id}_{\mathbf{R}}$$

et

$$f \circ g = \lim_{n \to +\infty} f_n \circ f_n^{-1} = \mathrm{Id}_{\mathbf{R}}.$$

□

## 2.2 Nombre de rotation de Poincaré

L'énoncé fondamental est le suivant :

THÉORÈME 2.2.1 : *Soit $f \in D^0(\mathbf{T}^1)$. Il existe $\rho \in \mathbf{R}$ tel que pour tout $x \in \mathbf{R}$ et tout $k \in \mathbf{Z}$, on a*

$$-1 < f^k(x) - x - k\rho < 1.$$

*En particulier, $\rho$ vérifie*

$$\rho = \lim_{k \mapsto \pm\infty} \frac{f^k(x)}{k},$$

*on l'appelle le* nombre de rotation *de $f$ et on le note $\rho(f)$. En d'autre termes, on a*

$$-1 < f^k(x) - T^k_{\rho(f)}(x) < 1$$

*pour tout $x \in \mathbf{R}$ et tout $k \in \mathbf{Z}$.*

*Preuve du théorème.* Commençons par écrire $f(x) = x + \varphi(x)$, où $\varphi : \mathbf{R} \to \mathbf{R}$ est périodique, de période 1.

LEMME 2.2.2 : *Pour tous réels $x$ et $y$, on a $-1 < \varphi(y) - \varphi(x) < 1$.*

*Preuve.* Soit $y'$ l'unique élément de $[x, x+1[$ tel que $y' - y \in \mathbf{Z}$. Des inégalités

$$x + \varphi(x) \leq y' + \varphi(y') < x + 1 + \varphi(x),$$

on déduit que

$$-1 = x - x - 1 < x - y' \leq \varphi(y') - \varphi(x) < x + 1 - y' \leq x + 1 - x = 1.$$

□

Si on applique le lemme précédent aux applications $f^k$, $k \geq 1$, on obtient

$$0 \leq M_k - m_k < 1,$$

où

$$m_k = \min_{x \in \mathbf{R}} f^k(x) - x, \;\; M_k = \max_{x \in \mathbf{R}} f^k(x) - x.$$

Pour tous entiers strictement positifs $k$ et $k'$, et pour tout réel $x$, on a

$$f^{k+k'}(x) - x = f^k(f^{k'}(x)) - f^{k'}(x) + f^{k'}(x) - x$$

et on a donc

$$m_k + m_{k'} \leq f^{k+k'}(x) - x \leq M_k + M_{k'},$$

ce qui implique

$$m_k + m_{k'} \leq m_{k+k'} \leq M_{k+k'} \leq M_k + M_{k'}.$$

Une récurrence évidente nous dit que pour tous entiers strictement positifs $k$ et $k'$, on a $M_{k'k} \leq k'M_k$ et $m_{k'k} \geq km_{k'}$ et donc que

$$\frac{m_{k'}}{k'} \leq \frac{m_{k'k}}{k'k} \leq \frac{M_{k'k}}{k'k} \leq \frac{M_k}{k}.$$

L'ensemble des nombres de la forme $\dfrac{m_k}{k}$ est donc à gauche de l'ensemble des nombres de la forme $\dfrac{M_k}{k}$. Puisque $\dfrac{M_k}{k} - \dfrac{m_k}{k} < \dfrac{1}{k}$, on en déduit que

$$\sup_{k\geq 1} \frac{m_k}{k} = \inf_{k\geq 1} \frac{M_k}{k}.$$

Si on note $\rho$ cette borne commune, on a $m_k \leq k\rho \leq M_k$ pour tout $k \geq 1$. Il existe $x_k$ et $y_k$ tels que $f^k(x_k) - x_k = m_k$ et $f^k(y_k) - y_k = M_k$. Le théorème des valeurs intermédiaires nous dit qu'il existe $z_k$ tel que $f^k(z_k) - z_k = k\rho$. Si on fixe $x \in \mathbf{R}$ et qu'on applique le lemme 2.2.2 à la fonction $f^k$ et aux points $x$ et $z_k$, on obtient

$$-1 < f^k(x) - x - k\rho < 1.$$

Le théorème 2.2.1 a donc été prouvé pour $k \geq 1$. Il est évident si $k = 0$ et il reste à le prouver si $k < 0$. Appliquons donc l'inégalité obtenue au point $f^k(x)$ pour l'itéré $f^{-k}$. On obtient

$$-1 < f^{-k}(f^k(x)) - f^k(x) + k\rho < 1,$$

ce qui implique

$$-1 < f^k(x) - x - k\rho < 1.$$

□

**Remarque** La preuve précédente nous donne les informations suivantes sur le nombre de rotation. Si $p \in \mathbf{Z}$ et $q \geq 1$ sont deux entiers, alors

**i)** $\rho(f) = p/q$ si et seulement s'il existe $x \in \mathbf{R}$ tel que $f^q(x) = x + p$ ;

**ii)** $\rho(f) > p/q$ si et seulement si $f^q(x) > x + p$ pour tout $x \in \mathbf{R}$ ;

**iii)** $\rho(f) < p/q$ si et seulement si $f^q(x) < x + p$ pour tout $x \in \mathbf{R}$.

Énonçons maintenant quelques propriétés du nombre de rotation.

Proposition 2.2.3 : **i)** *Le nombre de rotation de $T_a$ est $a$.*

**ii)** *Pour tout $f \in D^0(\mathbf{T}^1)$ et tout $p \in \mathbf{Z}$, on a $\rho(f + p) = \rho(f) + p$.*

**iii)** *Pour tout $f \in D^0(\mathbf{T}^1)$ et tout $q \in \mathbf{Z}$, on a $\rho(f^q) = q\rho(f)$.*

**iv)** *Si $f \leq g$, alors $\rho(f) \leq \rho(g)$.*

**v)** *L'application $f \mapsto \rho(f)$ est continue.*

*Preuve.* Pour obtenir **i)**, il suffit d'écrire

$$\rho(T_a) = \lim_{k\to+\infty} \frac{T_a^k(x)}{k} = \lim_{k\to+\infty} \frac{x+ka}{k} = a.$$

Pour montrer **ii)**, remarquons que

$$(f+p)^k = (T_p \circ f)^k = T_p^k \circ f^k$$

puisque $f$ et $T_p$ commutent. Ainsi, on a

$$\rho(f+p) = \lim_{k\to+\infty} \frac{(f+p)^k(x)}{k} = \lim_{k\to+\infty} \frac{f^k(x)+kp}{k} = \rho(f)+p.$$

Pour établir **iii)**, écrivons

$$\rho(f^q) = \lim_{k\to+\infty} \frac{(f^q)^k(x)}{k} = \lim_{k\to+\infty} \frac{f^{qk}(x)}{k} = q\rho(f).$$

Pour prouver **iv)**, commençons par montrer par récurrence que pour tout $k \geq 1$, on a $f^k \leq g^k$. L'inégalité est supposée vraie pour $k = 1$. Si elle est vraie au rang $k$ elle est également varie au rang $k+1$ car pour tout $x \in \mathbf{R}$ on a

$$f^{k+1}(x) = f(f^k(x)) \leq g(f^k(x)) \leq g(g^k(x)) = g^{k+1}(x).$$

Il reste à écrire

$$\rho(f) = \lim_{k\to+\infty} \frac{f^k(x)}{k} \leq \lim_{k\to+\infty} \frac{g^k(x)}{k} = \rho(g).$$

Il reste à montrer **v)**. Commençons par remarquer que les applications $f \mapsto f^k$ sont continues, puisque c'est le cas de l'application $(f,g) \to f \circ g$. Supposons donc que $\lim_{n\to+\infty} f_n = f$. Fixons $\varepsilon > 0$, puis choisissons $k \geq 1$ pour que $3 < k\varepsilon$. Fixons $x \in \mathbf{R}$. Il existe $N \geq 1$ tel que $-1 < f^k(x) - f_n^k(x) < 1$, pour tout $n \geq N$. Des inégalités

$$\begin{aligned}
-1 &< f_n^k(x) - x - k\rho(f_n) < 1, \\
-1 &< -f^k(x) + x + k\rho(f) < 1, \\
-1 &< f^k(x) - f_n^k(x) < 1,
\end{aligned}$$

on déduit

$$-3 < k(\rho(f) - \rho(f_n)) < 3,$$

ce qui implique

$$-\varepsilon < \rho(f) - \rho(f_n) < \varepsilon.$$

□

**Remarques**

1. Remarquons que $\rho(f^{-1}) = -\rho(f)$ et que $\rho(f^q + p) = q\rho(f) + p$, pour tous entiers $p$ et $q$.

2. L'assertion **ii)** implique que pour tout $F \in \mathrm{Homeo}_+(\mathbf{T}^1)$, la classe $\rho(f) + \mathbf{Z} \in \mathbf{T}^1$ ne dépend pas du choix du relèvement $f$ de $F$. C'est le *nombre de rotation* $\rho(F) \in \mathbf{T}^1$ de $F$.

3. La formule $\rho(f\circ g)=\rho(f)+\rho(g)$ est fausse en général (exercice **). Cependant cette formule est vraie si $f$ et $g$ commutent. En effet, des inégalités

$$\begin{aligned}-1 &< f^k\circ g^k(x)-g^k(x)-k\rho(f)<1,\\ -1 &< g^k(x)-x-k\rho(g)<1,\end{aligned}$$

on obtient

$$-2<(f\circ g)^k(x)-x-k(\rho(f)+\rho(g))<2,$$

ce qui implique

$$\rho(f\circ g)=\lim_{k\to+\infty}\frac{(f\circ g)^k(x)}{k}=\rho(f)+\rho(g).$$

**Exemple** Fixons $\alpha\in(-\frac{1}{2\pi},\frac{1}{2\pi})$ et définissons pour tout $t\in\mathbf{R}$ l'application

$$f_t\ :x\mapsto x+\alpha\sin(2\pi x)+t.$$

On vérifie que $f_t\in D^0(\mathbf{T}^1)$ et que $t\mapsto f_t$ est continue. La proposition 2.2.3 nous dit alors que l'application $r\ :t\mapsto\rho(f_t)$ est continue, croissante et vérifie $r(t+1)=r(t)+1$, pour tout $t\in\mathbf{R}$. On peut prouver que chaque intervalle $r^{-1}(\{a\})$ est réduit à un point si et seulement si $a\notin\mathbf{Q}$ (voir exercice**).

## 2.3 Dynamique des homéomorphismes de nombre de rotation rationnel

Commençons par étudier les homéomorphismes dont le nombre de rotation est nul.

PROPOSITION 2.3.1 : *Soit $f\in D^0(\mathbf{T}^1)$. Alors $\rho(f)=0$ si et seulement $f$ a un point fixe.*

*Preuve.* Si $f$ a un point fixe $x$, alors $\rho(f)=\lim_{k\to+\infty}\frac{f^k(x)}{k}=0$. Réciproquement, nous avons vu dans la section précédente que pour tout $f\in D^0(\mathbf{T}^1)$, il existe $x\in\mathbf{R}$ tel que $f(x)-x=\rho(f)$, ce qui implique la proposition. □

Donnons nous $f\in D^0(\mathbf{T}^1)$ tel que $\rho(f)=0$. L'ensemble $\mathrm{Fix}(f)$ est une partie fermée invariante par $T\ :x\mapsto x+1$. Si $]a,b[$ est une composante connexe de $\mathbf{R}\setminus\mathrm{Fix}(f)$, deux cas sont possibles suivants que la fonction $f-\mathrm{Id}_{\mathbf{R}}$ est positive ou négative sur $]a,b[$. Dans le premier cas, pour tout $x\in]a,b[$, la suite $(f^k(x))_{k\in\mathbf{Z}}$ est croissante et vérifie

$$\lim_{k\to-\infty}f^k(x)=a,\quad\lim_{k\to+\infty}f^k(x)=b\,;$$

dans le second cas, elle est décroissante et vérifie

$$\lim_{k\to-\infty}f^k(x)=b,\quad\lim_{k\to+\infty}f^k(x)=a.$$

Remarquons que les points fixes de l'homéomorphisme $F$ relevé par $f$ sont les images par la projection dans $\mathbf{T}^1$ des points fixes de $f$. Les ensembles $\alpha$ et $\omega$-limite d'un point $\widehat{x}\notin\mathrm{Fix}(F)$ sont les extrémités (égales si $F$ n'a qu'un point fixe !) de la composante connexe de $\mathbf{R}\setminus\mathrm{Fix}(F)$ qui contient $\widehat{x}$.

Passons maintenant au cas général. Nous allons voir qu'un homéomorphisme $F \in \mathrm{Homeo}_+(\mathbf{T}^1)$ a un nombre de rotation rationnel si et seulement s'il admet une orbite périodique et que dans ce cas toutes les périodes des orbites périodiques sont égales. Plus précisément :

PROPOSITION 2.3.2 : *Supposons que $F \in \mathrm{Homeo}_+(\mathbf{T}^1)$ a un nombre de rotation rationnel $\rho(F) = p/q + \mathbf{Z}$, où $p \in \mathbf{Z}$ et $q \geq 1$ sont premiers entre-eux. Alors*

**i)** *$F$ a une orbite périodique de période $q$ ;*

**ii)** *toutes les orbites périodiques de $F$ sont de période $q$ ;*

**iii)** *pour tout $\widehat{x} \in \mathbf{T}^1$, les ensembles $\alpha(\widehat{x})$ et $\omega(\widehat{x})$ sont des orbites périodiques.*

*Preuve.* Soit $f \in D^0(\mathbf{T}^1)$ le relèvement de $F$ tel que $\rho(f) = p/q$. Puisque $\rho(f^q - p) = 0$, on sait que $f^q - p$ a un point fixe. On en déduit que $F^q$ a un point fixe et que tout point fixe de $F^q$ est la projection d'un point fixe de $f^q - p$. Si $x + \mathbf{Z}$ est un point périodique de $F$ de période $q'$, il existe $p' \in \mathbf{Z}$ tel que $f^{q'}(x) = x + p'$. Remarquons que $f^{qq'}(x) - qp' = x$, ce qui implique que $0 = \rho(f^{qq'} - qp') = q'p - qp'$. Ainsi, il existe $r \geq 1$ tel que $q' = rq$ et $p' = rp$. On en déduit que $x$ est un point périodique de $f^q - p$ de période $r$. C'est donc un point fixe et on a $r = 1$. On vient d'établir **i)** et **ii)**. Pour prouver **iii)**, nous utilisons le fait que pour tout $\widehat{x} \in \mathbf{T}^1$, il existe un point fixe $\widehat{y}$ de $F^q$ tel que $\lim_{k \to +\infty} F^{kq}(\widehat{x}) = \widehat{y}$. C'est un point périodique de $F$, de période $q$, et son orbite est égale à $\omega(\widehat{x})$. On montre de façon similaire que $\alpha(\widehat{x})$ est une orbite périodique de $F$. □

## 2.4 Dynamique des homéomorphismes de nombre de rotation irrationnel

Le résultat principal est le résultat de semi-conjugaison suivant, dû à Poincaré :

THÉOREME 2.4.1 : *Soit $F \in \mathrm{Homeo}_+(\mathbf{T}^1)$. Si $\rho(F) \notin \mathbf{Q}/\mathbf{Z}$, alors, il existe une surjection continue $H : \mathbf{T}^1 \to \mathbf{T}^1$ de degré 1, relevée par une application croissante, telle que $H \circ F = T_{\rho(F)} \circ H$.*

*Preuve.* Fixons un relèvement $f \in D^0(\mathbf{T}^1)$ de $F$ et écrivons $\rho(f) = \rho$. Fixons également $x \in \mathbf{R}$. L'application

$$\begin{aligned} c\ &: \mathbf{Z}^2 \to \mathbf{R} \\ (p,q) &\mapsto q\rho - p \end{aligned}$$

est injective, puisque $\rho \notin \mathbf{Q}$, et son image est dense dans $\mathbf{R}$. L'application

$$\begin{aligned} c'\ &: \mathbf{Z}^2 \to \mathbf{R} \\ (p,q) &\mapsto f^q(x) - p \end{aligned}$$

est également injective. En effet

$$\begin{aligned} f^q(x) - p = f^{q'}(x) - p' &\Rightarrow f^q(x) - f^{q'}(x) = p - p' \\ &\Rightarrow f^{q-q'}(f^{q'}(x)) - f^{q'}(x) = p - p' \\ &\Rightarrow q = q' \text{ et } p = p', \end{aligned}$$

car $\rho \notin \mathbf{Q}$. Notons $Z$ l'image de $c$ et $Z'$ celle de $c'$. L'application $h = c \circ c'^{-1}$ est une bijection de $Z'$ sur $Z$. Notons que $Z$ et $Z'$ sont invariants par $T : x \mapsto x + 1$ et que $h \circ T = T \circ h$. Notons également que $C'$ est invariant par $f$, que $C$ est invariant par $T_\rho$ et que $h \circ f = T_\rho \circ h$.

Remarquons que $h^{-1}$ est croissant (et donc également $h$). En effet, si $q\rho - p < q'\rho - p'$, alors $(q-q')\rho < p - p'$. Dans le cas où $q = q'$, on en déduit $p > p'$ puis $f^q(x) - p < f^{q'}(x) - p'$. Dans le cas où $q > q'$, alors $\rho < \frac{p-p'}{q-q'}$ ce qui implique $f^{q-q'}(f^{q'}(x)) - f^{q'}(x) < p - p'$. Dans le cas où $q < q'$, alors $\rho > \frac{p-p'}{q-q'}$ ce qui implique $f^{q-q'}(f^{q'}(x)) - f^{q'}(x) > p - p'$.

L'image de $h$ étant dense dans $\mathbf{R}$, on peut étendre $h$ en une fonction définie sur $\mathbf{R}$ en posant

$$h(x) = \sup_{y \leq x,\, y \in Z'} h(y) = \inf_{y \geq x,\, y \in Z'} h(y).$$

On obtient une application croissante, dont l'image est dense et qui est donc continue et surjective. Elle vérifie

$$h(x+1) = \sup_{y \leq x+1,\, y \in Z'} h(y) = \sup_{y \leq x,\, y \in Z'} h(y) + 1 = h(x) + 1$$

et

$$h(f(x)) = \sup_{y \leq f(x),\, y \in Z'} h(y) = \sup_{y \leq x,\, y \in Z'} h(f(y)) = \sup_{y \leq x,\, y \in Z'} h(y) + \rho = h(x) + \rho.$$

La première égalité nous dit que $h$ relève une application $H\ : \mathbf{T}^1 \to \mathbf{T}^1$ de degré 1 ; la seconde nous dit que $h \circ f = T_\rho \circ h$ et donc que $H \circ F = T_{\rho(F)} \circ H$. $\square$.

**Remarques**

1. Remarquons que la préimage $h^{-1}(\{x\})$ de tout point $x \in \mathbf{R}$ est un intervalle compact $I$. L'ensemble des points $x$ tels que $h^{-1}(\{x\})$ n'est pas réduit à un point est une partie au plus dénombrable invariante par $T_{\rho(f)}$. On a des résultats similaires pour la préimage $H^{-1}(\{\widehat{x}\})$ d'un point $\widehat{x} \in \mathbf{T}^1$.

2. Si l'orbite de $x + \mathbf{Z}$ pour $F$ est dense, alors $Z$ est dense dans $\mathbf{R}$. Ceci implique que $h$ est un homéomorphisme Par conséquent $H$ est un homéomorphisme et $F$ est conjugué à $R_{\rho(F)}$. En d'autres termes, puisqu'un homéomorphisme de nombre de rotation rationnel n'est jamais transitif, nous en déduisons que pour un homéomorphisme du cercle $F$ préservant l'orientation, les propriétés suivantes sont équivalentes :

- $F$ est transitif ;
- $F$ est minimal ;
- $F$ est conjugué à une rotation d'angle irrationnel.

Continuons notre étude et expliquons plus précisément la dynamique des homéomorphismes $F \in \mathrm{Homeo}_+(\mathbf{T}^1)$ de nombre de rotation irrationnel, c'est-à-dire des homéomorphismes sans point périodique.

Théorème 2.4.2 : *Supposons que $F \in \mathrm{Homeo}_+(\mathbf{T}^1)$ n'a pas de point périodique. Il existe une partie fermée $X \subset \mathbf{T}^1$ vérifiant les propriétés suivantes :*

**i)** $X = \Omega(F)$ ;

**ii)** *$X$ est la seule partie minimale de $F$ ;*

**iii)** *pour tout $x \in \mathbf{T}^1$, on a $\alpha(x) = \omega(x) = X$ ;*

**iv)** *toute composante connexe de $\mathbf{T}^1 \setminus X$ est errante ;*

**v)** *si $X \neq \mathbf{T}^1$, alors $X$ est un ensemble de Cantor ;*

**vi)** *on a $X = \mathbf{T}^1$ si et seulement si $F$ est conjugué à une rotation irrationnelle.*

*Preuve.* Soit $X' \subset \mathbf{T}^1$ une partie fermée invariante non vide. Si $X' \neq \mathbf{T}^1$, son complémentaire est une union d'intervalles ouverts et $F$ induit une bijection $I \mapsto F(I)$ sur l'ensemble de ces intervalles. Aucun d'eux n'est périodique car $F$ n'a pas de point périodique et que les extrémités d'un intervalle périodique devraient être périodiques. On en déduit que toute composante connexe $I$ de $\mathbf{T}^1 \setminus X'$ est errante. On a donc $\Omega(F) \subset X'$. Comme conséquence, on sait que l'ensemble $X = \Omega(F)$ est contenu dans toute partie fermée invariante non vide. C'est donc un ensemble minimal, et même le seul. Pour tout point $x \in \mathbf{T}^1$, on sait que les ensembles $\alpha(x)$ et $\omega(x)$ sont inclus dans $\Omega(F)$, mais puisqu'ils contiennent également $X$, il coïncident avec $X$. Montrons maintenant que $X$ est un ensemble de Cantor si $X \neq \mathbf{T}^1$. Dans ce cas $\mathrm{Fr}(X)$ est une partie fermée invariante non vide. Elle contient donc $X$. Ceci signifie que $X$ est totalement discontinu. Puisque que $F$ n'a pas de point périodique, $X$ est infini et admet donc des points d'accumulation. L'ensemble des points d'accumulation est fermé et invariant, c'est donc l'ensemble $X$ tout entier : ce dernier ensemble n'a pas de point isolé. □

**Remarque** Dans le cas où l'application $H : \mathbf{T}^1 \to \mathbf{T}^1$ donnée par le théorème 4.4.1 est un homéomorphisme, on sait que $X = \mathbf{T}^1$. Dans le cas contraire, $X$ est l'ensemble des points $\widehat{x}$ tels que $H$ n'est pas constante au voisinage de $\widehat{x}$. En effet l'ensemble $X'$ des points vérifiant cette propriété, étant fermée et invariante par $F$, doit contenir $X$. Pour prouver l'inclusion inverse, remarquons que $H(X)$ est une partie fermée invariante par $T_\rho$, donc égale à $\mathbf{T}^1$. Par conséquent $H$ est constante sur chaque composante connexe de $\mathbf{T}^1 \setminus X$ (rappelons que $h$ est croissante) ce qui implique que $X$ contient $X'$. On peut être plus précis. Si on pose $I_{\widehat{y}} = H^{-1}(\{\widehat{y}\})$ pour tout $\widehat{y} \in \mathbf{T}^1$, on sait que $I_{T_{\widehat{\rho}}(\widehat{y})} = F(I_{\widehat{y}})$ puisque $H$ est une semi-conjugaison. Ceci implique que les intervalles $\mathrm{Int}(I_{\widehat{y}})$ sont errants, s'ils ne sont pas vides. L'orbite d'un tel intervalle est ordonnée cycliquement comme l'orbite d'un point par la rotation $T_{\widehat{\rho}}$. Ces intervalles sont exactement les composantes connexes de $\mathbf{T}^1 \setminus X$.

## 2.5 Difféomorphismes du cercle de nombre de rotation irrationnel

On peut se demander s'il est possible de construire des difféomorphismes du cercle sans point périodique qui ne sont pas conjugués à une rotation. De tels exemples, en classe $C^1$ on été donnés par Denjoy. Par contre, comme nous allons le voir, de tels exemples sont impossibles à construire en classe de différentiabilité plus grande. Pour tout $r \geq 1$, définissons le groupe $\mathrm{Diff}^r_+(\mathbf{T}^1)$ des difféomorphismes de classe $C^r$ de $\mathbf{T}^1$ préservant l'orientation. La dérivée $F'$ de $F \in \mathrm{Diff}^r_+(\mathbf{T}^1)$ est une application de $\mathbf{T}^1$ dans $\mathbf{R}$. On dira qu'une application $\psi : \mathbf{T}^1 \to \mathbf{R}$ est à *variation bornée* s'il existe $C > 0$ tel que pour toute famille $(\widehat{x}_i)_{i \in \mathbf{Z}/q\mathbf{Z}}$ cycliquement ordonnée sur le cercle, on a

$$\sum_{i \in \mathbf{Z}/q\mathbf{Z}} |\psi(\widehat{x}_{i+1}) - \psi(\widehat{x}_i)| \leq C.$$

C'est le cas, par exemple, si $\psi$ est de classe $C^1$ puisqu'on a

$$\sum_{i \in \mathbf{Z}/q\mathbf{Z}} |\psi(\widehat{x}_{i+1}) - \psi(\widehat{x}_i)| \leq \sum_{i \in \mathbf{Z}/q\mathbf{Z}} C\widehat{d}(\widehat{x}_i, \widehat{x}_{i+1}) \leq C,$$

où $C = \max_{\widehat{x} \in \mathbf{T}^1} |\psi'(\widehat{x})|$.

Le résultat suivant est dû à Denjoy :

THÉORÈME 2.5.1 : *Si la dérivée de $F \in \mathrm{Diffeo}^1_+(\mathbf{T}^1)$ est à variation bornée et si $F$ n'a pas de point périodique, alors $F$ est conjugué à une rotation d'angle irrationnel.*

*Preuve.* Remarquons d'abord que $\ln F'$ est à variation bornée. En effet, si on pose $m = \min_{\widehat{x}\in\mathbf{T}^1} F'(\widehat{x}) > 0$ et si on se donne une famille cycliquement ordonnée $(\widehat{x}_i)_{i\in\mathbf{Z}/q\mathbf{Z}}$, on a

$$\sum_{i\in\mathbf{Z}/q\mathbf{Z}} |\ln F'(\widehat{x}_{i+1}) - \ln F'(\widehat{x}_i)| \leq \frac{1}{m} \sum_{i\in\mathbf{Z}/q\mathbf{Z}} |F'(\widehat{x}_{i+1}) - F'(\widehat{x}_i)| \leq \frac{1}{m}\mathrm{Var}(F').$$

LEMME 2.5.2 : *Il existe une suite croissante $(q_n)_{n\geq 0}$ d'entiers positifs, telle que pour tout $n \geq 1$ et tout $\widehat{x} \in \mathbf{T}^1$ il existe un intervalle fermé $I_n$ de $\mathbf{T}^1$ joignant $\widehat{x}$ à $F^{q_n}(\widehat{x})$ dont les itérés $F^k(I_n)$, $0 \leq k \leq q_n$, sont disjoints deux à deux.*

*Preuve.* Puisque $F$ est semi-conjugué à une rotation d'angle irrationnel $R$ par une application qui préserve l'ordre cyclique, il suffit de prouver le lemme dans le cas d'une rotation $R$ et du point $\widehat{0} = 0 + \mathbf{Z}$. Écrivons alors $\widehat{x}_k = R^k(\widehat{0})$, si $k \in \mathbf{Z}$.

Posons $q_1 = 1$ et définissons par récurrence une suite $(q_n)_{n\geq 0}$ par la condition suivante

$$q_{n+1} = \inf\{q > q_n \mid d(\widehat{0}, \widehat{x}_q) < d(\widehat{0}, \widehat{x}_{q_n})\}.$$

Remarquons alors que

$$1 \leq q < q_n \Rightarrow d(\widehat{0}, \widehat{x}_{q_n}) < d(\widehat{0}, \widehat{x}_q).$$

Notons $I_n$ l'intervalle joignant $\widehat{0}$ à $\widehat{x}_{q_n}$ dont le diamètre est $d(\widehat{0}, \widehat{x}_{q_n})$. Supposons qu'il existe des entiers $q$ et $q'$ vérifiant $0 \leq q < q' \leq q_n$, tels que $R^q(I_n) \cap R^{q'}(I_n) \neq \emptyset$. On doit alors avoir $R^{q''}(I_n) \cap I_n \neq \emptyset$, où $q'' = q' - q$, ce qui est impossible. En effet, le point $x_{q''}$ n'appartient pas à $I_n$ car $d(\widehat{0}, \widehat{x}_{q''}) > d(\widehat{0}, \widehat{x}_{q_n})$. Pour les mêmes raisons, il n'appartient pas à $-I_{q_n}$, ce qui implique que $x_{q_n+q''}$ n'appartient pas à $I_{q_n}$. □

LEMME 2.5.3 : *Il existe une constante $C > 0$ telle que pour tout $\widehat{x} \in \mathbf{T}^1$, on a*

$$C^{-1} \leq (F^{q_n})'(\widehat{x})(F^{-q_n})'(\widehat{x}) \leq C.$$

*Preuve.* Remplaçant $\widehat{x}$ par $F^{q_n}(\widehat{x})$, on doit trouver une borne supérieure à

$$|\ln\left((F^{q_n})'(F^{q_n}(\widehat{x}))(F^{-q_n})'(F^{q_n}(\widehat{x}))\right)|.$$

On peut écrire cette quantité sous la forme

$$\left|\sum_{i=0}^{q_n-1} \ln F'(F^{q_n+i}(\widehat{x}) - \sum_{i=0}^{q_n-1} \ln F'(F^i(\widehat{x})\right|.$$

Elle est inférieure à

$$\sum_{i=0}^{q_n-1} |\ln F'(F^{q_n+i}(\widehat{x})) - \ln F'(F^i(\widehat{x}))|,$$

quantité qui admet $\mathrm{Var}\ln(F')$ comme majorant, d'après le lemme 4.5.2. On peut donc prendre

$$C = e^{\mathrm{Var}\ln(F')}.$$

□

*Preuve du théorème 2.5.1.* On doit prouver que $F$ n'a pas d'intervalle errant. Supposons que $I$ est un tel intervalle et notons $\mu$ la mesure de Lebesgue. On a

$$\lim_{n\to+\infty} \mu(F^{q_n}(I)) + \mu(F^{-q_n}(I)) = 0.$$

Or on sait que

$$\begin{aligned}
\mu(F^{q_n}(I)) + \mu(F^{-q_n}(I)) &= \int_I (F^{q_n})'\, d\mu + \int_I (F^{-q_n})'\, d\mu \\
&= \int_I (F^{q_n})' + (F^{-q_n})'\, d\mu \\
&\geq 2 \int_I \left((F^{q_n})'(F^{-q_n})'\right)^{1/2}\, d\mu \\
&\geq 2C^{-1/2}\mu(I).
\end{aligned}$$

□

## 3.1 Systèmes dynamiques mesurés

On va s'intéresser dans ce chapitre à un aspect différent des systèmes dynamiques. On se donnera un espace mesuré $(X, \mathcal{B}, \mu)$, où l'ensemble $X$ est muni d'une $\sigma$-algèbre $\mathcal{B}$ et d'une mesure $\mu$ sur $\mathcal{B}$. Si $T : X \to X$ est mesurable, rappelons que la mesure $T_*(\mu)$ est définie ainsi :

$$T_*(\mu)(A) = \mu(T^{-1}(A)) \text{ pour tout } A \in \mathcal{B},$$

et que c'est une mesure de probabilité si c'est le cas de $\mu$. Un système dynamique mesuré sera défini par une application mesurable $T : X \to X$ préservant $\mu$, c'est-à-dire telle que

$$\mu(A) = \mu(T^{-1}(A)) \text{ pour tout } A \in \mathcal{B}.$$

Si $T$ est inversible, on obtient un système dynamique inversible. Très souvent, l'inversibilité n'a lieu que presque partout, ce qui est suffisant pour avoir un système dynamique inversible : il existe une application mesurable $S : X \to X$, telle que $S \circ T(x) = T \circ S(x)$ pour presque tout point $x$. De façon équivalente, il existe $Y \in \mathcal{B}$ telle que $\mu(X \setminus Y) = 0$ et telle que $T$ induit une bijection mesurable en restriction à $Y$.

Il faut noter également que nous aurons très souvent une application $T : X \to X$ mesurable sur $\mathcal{B}$ (on parle alors de *système dynamique mesurable*) et qui laisse invariant plusieurs mesures (et définit donc plusieurs systèmes dynamiques mesurés).

On a également ici une notion de sous-système : on dira que $A \in \mathcal{B}$ est invariant si $T^{-1}(A) = A$. Ceci implique que

$$T|_A : (A, \mathcal{B}_A, \mu|_{\mathcal{B}_A}) \to (A, \mathcal{B}_A, \mu|_{\mathcal{B}_A})$$

est un système dynamique mesuré, où

$$\mathcal{B}_A = \{A \cap B\,,\, B \in \mathcal{B}\} = \{B \in \mathcal{B} \,|\, B \subset A\}.$$

Remarquons que la condition $T^{-1}(A) = A$ est plus forte que la condition $T(A) \subset A$ nécessaire pour définir un sous-système dynamique topologique. C'est la condition d'invariance adéquate dans le cas mesuré.

On a également une notion de facteur. Un système dynamique mesuré $(Y, \mathcal{N}, \nu, S)$ est un facteur de $(X, \mathcal{B}, \mu, T)$ s'il existe une application mesurable $H : X \to Y$ telle que $H \circ T = S \circ H$, presque partout et $H_*(\mu) = \nu$. Si $H$ est inversible (à ensembles négligeables près), les deux systèmes dynamiques sont conjugués.

**Exemples**

1. L'application $T : k \to k+1$ sur $\mathbf{Z}$ induit un système dynamique mesuré si $\mathbf{Z}$ est muni de mesure de comptage.

2. Si une application continue $T : X \to X$ sur un espace topologique a une orbite périodique $O$, la mesure de probabilité équidistribuée sur $O$ est une mesure borélienne préservée par $T$.

3. Les rotations $T_{\widehat{a}} : \mathbf{T}^r \to \mathbf{T}^r$ du tore définissent des systèmes dynamiques mesurés inversibles, si on munit $\mathbf{T}^r$ de la tribu borélienne et de la mesure de Haar.

4. Tout endomorphisme linéaire surjectif $\widehat{A}\ : \mathbf{T}^r \to \mathbf{T}^r$ définit également un système dynamique mesuré, si on munit $\mathbf{T}^r$ de la tribu borélienne et de la mesure de Haar.

5. Le décalage $\sigma\ : A^{\mathbf{N}} \to A^{\mathbf{N}}$ défini à partir d'un alphabet fini a beaucoup d'orbites périodiques et laisse donc invariant un grand nombre de mesures boréliennes de probabilité. Il admet d'autres mesures, plus intéressantes, en particulier les *mesures produits* et les *mesures de Markoff.*

Donnons-nous une famille $(p_a)_{a\in A}$ de nombres positifs tels que $\sum_{a\in A} p_a = 1$. Le théorème d'extension de Kolmogorov nous dit qu'il existe une unique mesure borélienne de probabilité $\mu$ sur $A^{\mathbf{N}}$ telle que pour tout cylindre $C_w^{n_0}$, où $w = (w_0, \ldots, w_{m-1})$, on a

$$\mu(C_w^{n_0}) = \prod_{0\leq n<m} p_{w_n}.$$

Remarquons que pour tout cylindre $C_w^{n_0}$, on a

$$\mu(\sigma^{-1}(C_w^{n_0})) = \mu(C_w^{n_0+1}) = \mu(C_w^{n_0}),$$

ce qui par unicité de l'extension, implique que $\sigma$ préserve $\mu$.

Un cas particulier est le cas où tous les $p_\alpha$ sont égaux, on parle alors de *mesure équidistribuée.* Le système dynamique est alors conjugué au système dynamique mesuré

$$\begin{aligned} F\ : (\mathbf{T}^1, \nu) &\to (\mathbf{T}^1, \nu) \\ \widehat{x} &\mapsto p\widehat{x} \end{aligned},$$

où $\nu$ est la mesure de Haar et $p = \sharp A$. En effet, on peut supposer que $A = \{0, \ldots, p-1\}$ et considérer la semi-conjugaison (au sens topologique)

$$\begin{aligned} H\ : \{0, \ldots, p-1\}^{\mathbf{N}} &\to \mathbf{T}^1 \\ x &\mapsto \sum_{n=0}^{\infty} \frac{x_n}{p^{n+1}} + \mathbf{Z} \end{aligned}.$$

On vérifie aisément que $\nu(H(C_w^{n_0})) = \mu(C_w^{n_0})$ pour tout cylindre $C_w^{n_0}$. Or, en dehors des points $p$-adiques qui ont deux antécédents et qui sont en nombre dénombrable, tous les autres points ont un unique antécédent. Le complémentaire $X$ des points $p$-adiques est de mesure de Haar totale et vérifie $F^{-1}(X) = X$. La réduction de $H$ à $H^{-1}(X)$ est une bijection de $H^{-1}(X)$ sur $X$

Introduisons maintenant la classe plus générale des mesures de Markov. On supposera pour faciliter les notations que $A = \{1, \ldots, p\}$. Soit $M = (M_{i,j})_{i,j}$ une matrice carrée d'ordre $p$ qui est *stochastique*, c'est-à-dire une matrice à coefficients positifs vérifiant $\sum_{j=1}^{p} M_{i,j} = 1$, pour tout $i \in \{1, \ldots, p\}$. Commençons par montrer qu'il existe au moins un vecteur $v = (v_1, \ldots, v_p)$ à coefficients positifs tels que $\sum_{i=1}^{p} v_i = 1$ et qui vérifie $vM = v$. On cherche ce vecteur comme point fixe de l'endomorphisme $L$ dont $M^t$ est la trace dans la base canonique. Remarquons que $L$ laisse invariant l'hyperplan affine $H$ d'équation $x_1 + \ldots + x_p = 1$. Il envoie la partie compacte convexe

$$C = \{(x_1, \ldots, x_p) \in H \,|x_i \geq 0, \ \text{pour tout}\ \ i \in \{1, \ldots, p\}\}$$

dans elle même. Fixons $w \in C$ et définissons

$$w_n = \frac{1}{n}\sum_{k=0}^{n-1} L^k(w).$$

Remarquons que

$$\|L(w_n) - w_n\| = \left\| \frac{1}{n}(L^n(w) - w) \right\| \leq \frac{1}{n} \text{diam}(C),$$

ce qui implique que toute valeur d'adhérence $v$ de la suite $(w_n)_{n\geq 1}$ est fixe par $L$. Ce vecteur est d'ailleurs unique dans le cas où il existe une puissance de $M$ dont les coefficients sont strictement positifs (c'est un cas particulier du théorème de Perron-Frobenius). En effet si $v$ et $v'$ sont deux vecteurs dans $C$ qui sont fixes par $L$, la droite affine passant par $v$ et $v'$ est formée de points fixes et contient un point de la frontière de $C$. Mais si les coefficients de $M^n$ sont strictement positifs, l'image par $L^n$ de tout point de $C$ est inclus dans l'intérieur (relatif) de $C$. Montrons maintenant que $\bigcap_{n\geq 0} L^n(C) = \{v\}$, ce qui impliquera que pour tout $w \in C$, on a $\lim_{n\to+\infty} L^n(w) = v$ et donc en écrivant $v = (v_1, \ldots, v_p)$ et en notant $(M^n_{i,j})_{i,j}$ la matrice $M^n$, que l'on a $\lim_{n\to+\infty} M^n_{i,j} = v_j$. Chaque ensemble $L^n(C)$ étant un simplexe affine de dimension $< p$, l'ensemble $\bigcap_{n\geq 0} L^n(C)$ est également un simplexe affine de dimension $q < p$, c'est l'ensemble des barycentres à coefficients positifs de $q + 1$ points. Puisque $\bigcap_{n\geq 0} L^n(C)$ est invariant par $L$, les $q + 1$ points extremaux sont tous fixes pour une même puissance de $L$. Ceci est impossible si $q \neq 0$ car on aurait un point périodique de $L$ sur la frontière de $C$. Si $v$ vérifie $vM = M$, il existe alors une unique mesure borélienne de probabilité $\mu$ telle que pour tout $w = (w_0, \ldots, w_{m-1}) \in \{1, \ldots, p\}^m$ et tout $n_0 \in \mathbf{N}$, on a

$$\mu(C^{n_0}_w) = v_{w_0} \left( \prod_{k=0}^{m-2} M_{w_k, w_{k+1}} \right).$$

En effet si on écrit

$$iw = (i, w_0, \ldots, w_{m-1}), \;\; wi = (w_0, \ldots, w_{m-1}, i),$$

pour $w = (w_0, \ldots, w_{m-1}) \in \{1, \ldots, p\}^m$, et $i \in \{1, \ldots, p\}$, la formule précédente implique que pour tout $n_0 \in \mathbf{N}$, on a

$$\sum_{i=1}^{p} \mu(C^{n_0}_i) = 1$$

et que pour tout $w = (w_0, \ldots, w_{m-1}) \in \{1, \ldots, p\}^{m+1}$ et tout $n_0 \in \mathbf{N}$, on a

$$\mu(C^{n_0}_w) = \sum_{i=1}^{p} \mu(C^{n_0-1}_{iw}) = \sum_{i=1}^{p} \mu(C^{n_0}_{wi})$$

(la première égalité de la ligne précédente est due au fait que $vM = v$, la seconde au fait que $M$ est une matrice stochastique). Le théorème d'extension de Caratheodory nous dit qu'il existe alors une unique mesure borélienne de probabilité $\mu$ définie par ces égalités. Là encore,on a

$$\mu(\sigma^{-1}(C^{n_0}_w)) = \mu(C^{n_0+1}_w) = \mu(C^{n_0}_w),$$

ce qui prouve que $\mu$ est invariante par $\sigma$.

6. On construit de façon équivalente des mesures produits ou des mesures de Markov sur $A^{\mathbf{Z}}$ qui sont invariantes par le décalage bilatéral.

7. Concluons cette section par un dernier exemple, propre au cas mesuré. Définissons sur $[0, 1[$ l'application

$$T \;: x \mapsto \left\{ \frac{1}{x} \right\} = \frac{1}{x} - \left[ \frac{1}{x} \right],$$

où $\{y\}$ est la partie fractionnaire et $[y]$ la partie entière d'un réel $y$. Elle n'est pas définie en 0. On a donc

$$x = \frac{1}{a(x) + T(x)},$$

où $a(x) = [1/x]$. Pour tout $k \geq 1$, l'ensemble $a^{-1}(\{k\})$ coïncide avec l'intervalle $]1/(k+1), 1/k]$. L'application $T$ induit une bijection strictement décroissante de $]1/(k+1), 1/k]$ sur $[0, 1[$. Si on fixe $a_1 \geq 1, \ldots, a_n \geq 1$, on en déduit que l'ensemble $I_{a_1,\ldots,a_n}$ des points $x \in [0, 1[$ tels que les itérés $T(x), \ldots, T^{n-1}(x)$, sont bien définis, tels que $T^{n-1}(x) \neq 0$, et tels que $a(x) = a_1, \ldots, a(T^{n-1}(x)) = a_n$, est un intervalle semi-ouvert et que l'application $T^n$ induit une bijection de $I_{a_1,\ldots,a_n}$ sur $[0, 1[$, croissante si $n$ est pair et décroissante si $n$ est impair. Le réciproque de cette bijection s'exprime explicitement. Remarquons que pour tout $x \in I_{a_1,\ldots,a_n}$, on a

$$x = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\frac{\cdots}{a_{n-1} + \frac{1}{a_n + T^n(x)}}}}},$$

que l'on peut écrire

$$x = \frac{p_n + p_{n-1}T^n(x)}{q_n + q_{n-1}T^n(x)},$$

où les deux suites $(p_k)_{k \geq n}$ et $(q_k)_{k \geq n}$ sont définies par les égalités

$$\begin{pmatrix} p_1 & q_1 \\ p_0 & q_0 \end{pmatrix} = \begin{pmatrix} 1 & a_1 \\ 0 & 1 \end{pmatrix}$$

et

$$\begin{cases} p_k = a_k p_{k-1} + p_{k-2} \\ q_k = a_k q_{k-1} + q_{k-2} \end{cases}$$

On peut vérifier que la suite $(q_n)_{n \geq 0}$ est strictement croissante et que

$$p_k q_{k-1} - p_{k-1} q_k = (-1)^{k-1}.$$

Les extrémités de $I_{a_1,\ldots,a_n}$ étant $p_n/q_n$ et $(p_n + p_{n-1})/(q_n + q_{n-1})$, on en déduit que la longueur de $I_{a_1,\ldots,a_n}$ vérifie

$$\operatorname{long}(I_{a_1,\ldots,a_n}) \leq \frac{1}{q_n(q_n + q_{n-1})} \leq \frac{1}{q_n^2} \leq \frac{1}{(n+1)^2}.$$

Pour tout point $x$, la suite $(a_1(T^n(x))_{n \geq 0}$ détermine le développement en fraction continue de $x$. Il n'est pas difficile de voir que la suite des itérés $T^n(x)$ aboutit en 0 si et seulement $x$ est rationnel. En d'autres termes la suite $(T^n(x))_{n \geq 0}$ est bien définie si et seulement si $x$ est irrationnel et on a un système dynamique bien déterminé défini sur $[0, 1[\setminus \mathbf{Q}$ en prenant la restriction de $T$ à cet ensemble. Puisque pour toute suite $(a_i)_{i \geq 1}$ d'entiers strictement positifs, et pour tout $n \geq 0$, on a

$$\operatorname{long}(I_{a_1,\ldots,a_n}) \leq \frac{1}{(n+1)^2},$$

on en déduit que l'application

$$\begin{aligned} H \ : [0, 1[\setminus \mathbf{Q} &\to (\mathbf{N} \setminus \{0\})^{\mathbf{N} \setminus \{0\}} \\ x &\mapsto (a(T^{i-1}(x)) \end{aligned}$$

conjugue $T|_{[0,1[\setminus \mathbf{Q}}$ au décalage $(a_n)_{n \geq 1} \mapsto (a_{n+1})_{n \geq 1}$ sur $(\mathbf{N} \setminus \{0\})^{\mathbf{N} \setminus \{0\}}$.

Remarquons maintenant que $T$ préserve la *mesure de Gauss*

$$d\mu = \frac{1}{\ln 2} \frac{dx}{1+x}.$$

Il suffit de vérifier que pour tous intervalle $[a,b] \subset [0,1[$, on a $\mu(T^{-1}([a,b]) = \mu([a,b])$. Or on a

$$\begin{aligned}
\ln 2\, \mu(T^{-1}([a,b])) &= \sum_{k=1}^{\infty} \int_{\frac{1}{b+k}}^{\frac{1}{a+k}} \frac{1}{1+t} dt \\
&= \sum_{k=1}^{\infty} \ln(1+\frac{1}{a+k}) - \ln(1+\frac{1}{b+k}) \\
&= \sum_{k=1}^{\infty} \ln(a+k+1) - \ln(a+k) - \ln(b+k+1) + \ln(b+k) \\
&= \ln(b+1) - \ln(a+1) = \ln 2\, \mu([a,b]).
\end{aligned}$$

## 3.2 Théorèmes asymptotiques

Commençons par le *théorème de récurrence de Poincaré.*

THÉORÈME 3.2.1 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) < \infty$. Si $A \in \mathcal{B}$ vérifie $\mu(A) > 0$, alors l'ensemble $B$ des points $x \in A$ tels qu'il existe une infinité d'entiers $n \geq 1$ pour lesquels $T^n(x) \in A$ est une partie mesurable qui vérifie $\mu(B) = \mu(A)$.*

*Preuve.* Le complémentaire de $B$ dans $A$ s'écrit $A \setminus B = \bigcup_{n\geq 1} E_n$, où

$$E_n = \{x \in A \mid m \geq n \Rightarrow T^m(x) \notin A\}.$$

Remarquons que les ensembles $T^{-mn}(E_n)$, $m \geq 1$ sont tous disjoints de $E_n$. Ceci implique que les ensembles $T^{-mn}(E_n)$, $m \geq 0$, sont disjoints deux à deux. Puisqu'ils ont même mesure et puisque $\mu(X) < +\infty$, cette mesure commune est nulle. Ainsi, on a $\mu\left(\bigcup_{n\geq 1} E_n\right) = 0$. □.

COROLLAIRE 3.2.2 : *Soit $T : X \to X$ une transformation continue définie sur un espace métrique $(X,d)$. Supposons qu'il existe une mesure borélienne finie invariante par $T$ dont le support est $X$. Alors, il n'y a pas de point errant. De plus, si $X$ est séparable, presque tout point est positivement récurrent.*

*Preuve.* Le support de la mesure est le plus petit ensemble fermé dont le complémentaire est de mesure nulle. Dire que c'est $X$ signifie que tout ensemble ouvert non vide est de mesure strictement positive. Le théorème 3.2.1 nous dit donc qu'il n'y a pas d'ensemble ouvert errant. Pour montrer le second point, considérons une base dénombrable $(U_i)_{i\in I}$ de la topologie de $X$ et pour tout $i \in I$, notons $E_i$ l'ensemble des points $x \in U_i$ dont l'orbite (strictement) positive est disjointe de $U_i$. D'après le théorème 3.2.1, on sait que $\mu(E_i) = 0$ et donc que $\mu\left(\bigcup_{i\in I} E_i\right) = 0$. Le complémentaire ce ce dernier ensemble n'est rien d'autre que l'ensemble des points positivement récurrents. □

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré. Si $f = \sum_{1 \leq i \leq n} \lambda_i \chi_{A_i}$ est une fonction étagée, où $\lambda_i \in \mathbf{C}$, $A_i \in \mathcal{B}$ et $\mu(A_i) < +\infty$, alors pour tout $p \geq 1$, on a

$$\begin{aligned} \int f^p \circ T d\mu &= \sum_{i=1}^{n} \lambda_i^p \int \chi_{A_i} \circ T \, d\mu \\ &= \sum_{i=1}^{n} \lambda_i^p \int \chi_{T^{-1}(A_i)} \, d\mu \\ &= \sum_{i=1}^{n} \lambda_i^p \mu(T^{-1}(A_i)) \, d\mu \\ &= \sum_{i=1}^{n} \lambda_i^p \mu(A_i) \, d\mu = \int f^p \, d\mu \end{aligned} .$$

L'ensemble des fonctions étagées étant dense dans $L^p(\mu)$, on en déduit que pour tout $f \in L^p(\mu)$, l'application $f \circ T$ appartient à $L^p(\mu)$ et vérifie $\int f^p \circ T \, d\mu = \int f^p \, d\mu$. On obtient ainsi une autre façon de définir l'invariance de la mesure. Un cas particulièrement intéressant est celui où $p = 2$. L'opérateur

$$\begin{aligned} U_T \ : L^2(\mu) &\to L^2(\mu) \\ f &\mapsto f \circ T \end{aligned}$$

est alors un opérateur unitaire de $L^2(\mu)$ : on a

$$\langle U_T(f), U_T(g) \rangle = \int (f \circ T)(\overline{g} \circ T) \, d\mu = \int f \overline{g} \, d\mu = \langle f, g \rangle.$$

Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré. Les *moyennes de Birkhoff* de $f \ : X \to \mathbf{C}$ sont les moyennes temporelles $\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i$. On veut étudier ces moyennes. Commençons par énoncer le théorème de Von Neumann qui donne une convergence, au sens quadratique :

Théorème 3.2.3 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré. Définissons l'espace $L^2_{\text{inv}}(\mu) = \text{Fix}(U_T)$ ainsi que la projection orthogonale $p_{\text{inv}} \ : L^2(\mu) \to L^2_{\text{inv}}(\mu)$. Alors, pour tout $f \in L^2(\mu)$, on a*

$$p_{\text{inv}}(f) = \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} U_T^i(f).$$

*Preuve.* On doit prouver que

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} U_T^i(f) = f,$$

si $f \in L^2_{\text{inv}}(\mu)$ (ce qui évident) et

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} U_T^i(f) = 0$$

si $f \in L^2_{\text{inv}}(\mu)^\perp$ (ce qui l'est moins).

Commençons par prouver que

$$L^2_{\text{inv}}(\mu)^\perp = \overline{\text{Im}(\text{Id} - U_T)}$$

ou, de façon équivalente, que

$$L^2_{\mathrm{inv}}(\mu) = \mathrm{Im}(\mathrm{Id} - U_T)^{\perp}.$$

Une fonction $f$ appartient à $(\mathrm{Id} - U_T)^{\perp}$ si et seulement si, pour tout $g \in L^2(\mu)$, on a

$$\langle f, g - U_T(g) \rangle = \langle f - U_T^*(f), g \rangle = 0.$$

Ainsi $\mathrm{Im}(\mathrm{Id} - U_T)^{\perp} = \mathrm{Fix}(U_T^*)$. Pour prouver que $\mathrm{Fix}(U_T) = \mathrm{Fix}(U_T^*)$, rappelons que

$$\|U_T^*\| = \sup_{\|f\| \leq 1,\, \|g\| \leq 1} |\langle U_T^*(f), g \rangle| = \sup_{\|f\| \leq 1,\, \|g\| \leq 1} |\langle f, U_T(g) \rangle| = \|U_T\|.$$

Ceci implique que pour tout $f \in L^2_{\mathrm{inv}} = \mathrm{Fix}(U_T)$, on a

$$\begin{aligned}\langle U_T^*(f) - f, U_T^*(f) - f \rangle &= \langle U_T^*(f), U_T^*(f) \rangle - \langle f, U_T^*(f) \rangle - \langle U_T^*(f), f \rangle + \langle f, f \rangle \\ &\leq \langle f, f \rangle - \langle U_T(f), f \rangle - \langle f, U_T(f) \rangle + \langle f, f) \rangle = 0\end{aligned},$$

et donc $f \in \mathrm{Fix}(U_T^*)$. Similairement, pour tout $f \in \mathrm{Fix}(U_T^*)$, on a

$$\begin{aligned}\langle U_T(f) - f, U_T(f) - f \rangle &= \langle U_T(f), U_T(f) \rangle - \langle f, U_T(f) \rangle - \langle U_T(f), f \rangle + \langle f, f \rangle \\ &= \langle f, f \rangle - \langle U_T^*(f), f \rangle - \langle f, U_T s^*(f) \rangle + \langle f, f) \rangle = 0\end{aligned},$$

et donc $f \in \mathrm{Fix}(U_T)$.

Supposons donc que $f \in L^2_{\mathrm{inv}}(\mu)^{\perp}$. Pour tout $\varepsilon > 0$, on peut trouver $g \in L^2(\mu)$ tel que $\|f - (U_T(g) - g)\| \leq \varepsilon$. Ceci implique que

$$\frac{1}{n} \left\| \sum_{i=0}^{n-1} U_T^i(f) - U_T^n(g) - g \right\| \leq \varepsilon,$$

et donc que

$$\frac{1}{n} \left\| \sum_{i=0}^{n-1} U_T^i(f) \right\| \leq \frac{1}{n}(\|U_T^n(g)\| + \|g\|) + \varepsilon = \frac{2}{n}\|g\| + \varepsilon \leq 3\varepsilon,$$

si $n$ est grand. □

Énonçons maintenant le *théorème ergodique de Birkhoff* qui donne un résultat de convergence presque sure.

THÉORÈME 3.2.4 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré. Fixons $f \in L^1(\mu)$. Il existe alors une fonction $f^* \in L^1(\mu)$, invariante by $T$, vérifiant :*

**i)** *pour presque tout point $x \in X$, on a $\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x) = f^*(x)$ ;*

**ii)** *on a $\|f^*\|_{L^1} \leq \|f\|_{L^1}$ ;*

**iii)** *si $\mu(X) < +\infty$, alors $\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i = f^*$ dans $L^1(\mu)$ et par conséquent $\int f^*\, d\mu = \int f\, d\mu$ ;*

**iv)** *si $T$ est inversible, alors pour presque tout point $x \in X$, on a $\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^{-i}(x) = f^*(x)$.*

*Preuve.* Considérons $f \in L^1(\mu)$ et pour tout $n \geq 1$, posons

$$S_n f = \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i.$$

Commençons par prouver que la suite $(S_n f(x))_{n\geq 0}$ converge presque surement dans $\mathbf{R}\cup\{-\infty,+\infty\}$. Posons

$$\overline{f}(x) = \limsup_{n\to+\infty} S_n f(x), \;\; \underline{f}(x) = \liminf_{n\to+\infty} S_n f(x).$$

Remarquons que $\overline{f}$ et $\underline{f}$ sont invariantes car

$$S_{n+1}f = \frac{1}{n+1} f + \frac{n}{n+1} S_n f \circ T.$$

Commençons par prouver que si $\alpha < \beta$, alors $\mu(E_{\alpha,\beta}) = 0$, où

$$E_{\alpha,\beta} = \{x \in X \;\mid\; \underline{f}(x) < \alpha < \beta < \overline{f}(x)\}.$$

Ceci impliquera que $\overline{f}(x) = \underline{f}(x)$, pour presque tout $x$, puisque

$$\{x \in X \;\mid\; \underline{f}(x) < \overline{f}(x)\} = \bigcup_{\alpha,\beta\in\mathbf{Q},\alpha<\beta} E_{\alpha,\beta}.$$

Nous allons utiliser le *lemme ergodique maximal.*

Lemme 3.2.5 : *On a*

$$\int_{\widetilde{f}(x)>0} f\,d\mu \geq 0,$$

*où on pose*

$$\widetilde{f}(x) = \sup_{n\geq 1} \sum_{i=0}^{n-1} f(T^i(x)).$$

*Preuve.* Pour tout $n \geq 1$ et tout $x \in X$, on définit

$$f_n(x) = \max(0, f(x), f(x) + f(T(x)), \ldots, f(x) + f(T(x)) + \ldots + f(T^{n-1}(x))),$$

et

$$E_n = \{x \in X \;\mid\; f_n(x) > 0\}.$$

Dans le cas où $x \in E_n$, on a $f_n(x) = f(x) + f_{n-1}(T(x))$; dans le cas où $x \notin E_n$ on a $f_n(x) = 0$ et $f_{n-1}(T(x)) \geq 0$. Ainsi,

$$\begin{aligned}\int_{E_n} f\,d\mu &= \int_{E_n} f_n - f_{n-1}\circ T\,d\mu \\ &\geq \int_X f_n - f_{n-1}\circ T\,d\mu \\ &= \int_X f_n - f_{n-1}\,d\mu \geq 0,\end{aligned}$$

car $f_n \geq f_{n-1}$. Puisque

$$\{x \in X \;\mid\; \widetilde{f}(x) > 0\} = \bigcup_{n\geq 1} E_n,$$

on obtient

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

□

Montrons maintenant que $\mu(E_{\alpha,\beta}) = 0$ si $\mu(E_{\alpha,\beta}) < \infty$. Pour cela, appliquons le lemme 5.2.5 à la restriction $T|_{E_{\alpha,\beta}}$ (remarquons que $E_{\alpha,\beta}$ est invariant) et à l'application intégrable $f - \beta$. On obtient

$$\int_{E_{\alpha,\beta}} f - \beta \, d\mu \geq 0.$$

Appliquons-la également à $\alpha - f$. On obtient

$$\int_{E_{\alpha,\beta}} \alpha - f \, d\mu \geq 0.$$

Finalement, on obtient

$$\alpha\mu(E_{\alpha,\beta}) \geq \int_{E_{\alpha,\beta}} f \, d\mu \geq \beta\mu(E_{\alpha,\beta})$$

ce qui implique que $\mu(E_{\alpha,\beta}) = 0$.

Il reste à prouver que $\mu(E_{\alpha,\beta}) < \infty$. Nous établirons pour cela que tout ensemble $C \subset E_{\alpha,\beta}$ de mesure finie vérifie :

$$\begin{cases} \mu(C) \leq \frac{1}{\beta} \int_{E_{\alpha,\beta}} f \, d\mu & \text{si } \beta > 0 \\ \mu(C) \leq \frac{1}{\alpha} \int_{E_{\alpha,\beta}} f \, d\mu & \text{si } \alpha < 0. \end{cases}$$

Dans le cas où $\beta > 0$, on applique le lemme ergodique maximal à la restriction $T|_{E_{\alpha,\beta}}$ et à l'application intégrable $f - \beta\chi_C$. Puisque $f - \beta\chi_C \geq f - \beta$, on obtient

$$\int_{E_{\alpha,\beta}} f - \beta\chi_C \, d\mu \geq 0$$

et donc

$$\beta\mu(C) \leq \int_{E_{\alpha,\beta}} f \, d\mu.$$

Dans le cas où $\alpha < 0$, on applique le lemme à l'application $\alpha\chi_C - f$. Puisque $\alpha\chi_C - f \geq \alpha - f$, on obtient

$$\int_{E_{\alpha,\beta}} \alpha\chi_C - f \, d\mu \geq 0$$

et donc

$$\alpha\mu(C) \geq \int_{E_{\alpha,\beta}} f \, d\mu.$$

On va maintenant prouver **ii)**. Remarquons que

$$\int |S_n f| \, d\mu \leq \frac{1}{n} \sum_{i=0}^{n-1} \int |f \circ T^i| \, d\mu = \int |f| \, d\mu,$$

ce qui, grâce au lemme de Fatou, implique

$$\int |f^*| \, d\mu = \int \liminf_{n \to +\infty} |S_n f| \, d\mu \leq \liminf_{n \to +\infty} \int |S_n f| \, d\mu \leq \int |f| \, d\mu.$$

Pour obtenir **iii)**, il faut montrer que

$$\lim_{n \to +\infty} \int |S_n f - f^*| \, d\mu = 0,$$

si $\mu(X) < +\infty$. Dans le cas où $f$ est bornée (disons par $M$), c'est une conséquence immédiate du théorème de convergence dominée de Lebesgue puisque $|S_n f(x) - f^*(x)| \to 0$ pour presque tout point $x$ et $|S_n f(x) - f^*(x)| \leq 2M$. Dans le cas où $f$ n'est pas bornée, alors pour tout $\varepsilon > 0$, on peut trouver $g \in L^\infty$ telle que $\int |f - g| \, d\mu < \varepsilon/3$. Ceci implique que pour tout $n \geq 1$, on a $\int |S_n f - S_n g| \, d\mu < \varepsilon/3$ et également que $\int |f^* - g^*| \, d\mu < \varepsilon/3$. Si $n$ est assez grand, on aura $\int |S_n g - g^*| \, d\mu < \varepsilon/3$, ce qui implique $\int |S_n f - f^*| \, d\mu < \varepsilon$.

Il reste à prouver **iv)**. Appliquons ce qui vient d'être fait à $T^{-1}$. On sait que la suite

$$S_{-n} f = \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^{-i}$$

converge presque sûrement vers une fonction $f^{**}$. On veut prouver que $f^* = f^{**}$ presque partout. Il suffit, bien sûr, de prouver que $f^*(x) \leq f^{**}(x)$ pour presque tout $x$ et donc de prouver que si $\alpha < \beta$, alors $\mu(F_{\alpha,\beta}) = 0$, où

$$F_{\alpha,\beta} = \{x \in X \mid f^{**}(x) < \alpha < \beta < f^*(x)\}.$$

Grâce au lemme ergodique maximal, on prouve, comme un peu plus haut, que pour toute partie $C \subset F_{\alpha,\beta}$ de mesure finie, on a

$$\begin{cases} \mu(C) \leq \frac{1}{\beta} \int_{F_{\alpha,\beta}} f \, d\mu & \text{si } \beta > 0 \\ \mu(C) \leq \frac{1}{\alpha} \int_{F_{\alpha,\beta}} f \, d\mu & \text{si } \alpha < 0. \end{cases}$$

On en déduit que $\mu(F_{\alpha,\beta}) < +\infty$, puis, grâce encore au lemme ergodique maximal, que

$$\alpha \mu(F_{\alpha,\beta}) \geq \int_{F_{\alpha,\beta}} f \, d\mu \geq \beta \mu(F_{\alpha,\beta}).$$

□

## 3.3 Propriétés des systèmes dynamiques mesurés

Comme dans le cas topologique, on va introduire plusieurs définitions sur les systèmes dynamiques mesurés. Commençons par un résultat d'indécomposabilité analogue à la transitivité.

PROPOSITION 3.3.1 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré. Les deux propriétés suivantes sont équivalentes :*

**i)** *si $A \in \mathcal{B}$ vérifie $T^{-1}(A) = A$, alors $\mu(A) = 0$ ou $\mu(X \setminus A) = 0$ ;*

**ii)** *toute fonction mesurable $f : X \to \mathbf{C}$ invariante par $T$ est constante presque partout.*

*Si ces propriétés sont vérifiées, on dit que le système est* ergodique.

*Preuve.* Supposons que **ii)** est vraie. Si $A \in \mathcal{B}$ vérifie $T^{-1}(A) = A$, alors la fonction $\chi_A$ est invariante. Puisqu'elle constante presque surement, l'un des ensembles $A$ ou $X \setminus A$ est de mesure nulle.

Supposons au contraire que **i)** est vraie. Si $f : X \to \mathbf{C}$ est invariante par $T$, alors pour toute partie boréliennne $B \subset \mathbf{C}$, l'ensemble $A = f^{-1}(B)$ est mesurable et vérifie $T^{-1}(A) = A$. Ainsi, on a $\mu(A) = 0$ ou $\mu(X \setminus A) = 0$. Écrivons $f = f_0 + i f_1$ où $f_0$ et $f_1$ sont à valeurs réelles. On en déduit que pour tout $n \geq 1$, il existe un unique $k_n \in \mathbf{Z}$ tel que $f_0(x) \in [k_n/n, (k_n+1)/n[$ presque surement. On en déduit que $f_0$ (ainsi bien sûr que $f_1$) est constante presque surement. □

**Remarque** Dans le cas où $T$ admet plusieurs mesures invariantes et où le système dynamique $(X, \mathcal{B}, \mu, T)$ est ergodique, on dira plus simplement que $\mu$ est ergodique.

Dans le cas où la mesure est finie, on a le critère d'ergodicité évident (mais utile) suivant :

Proposition 3.3.2 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré, où $\mu(X) < +\infty$. Fixons $p \geq 1$. Les deux propriétés suivantes sont équivalentes :*

**i)** *le système est ergodique ;*

**ii)** *toute fonction $f \in L^p(\mu)$ qui est invariante est constante presque partout.*

Le résultat suivant nous dit que pour un système ergodique, les moyennes temporelles et spatiales coïncident (dans le cas où la mesure invariante est finie) :

Proposition 3.3.3 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré ergodique et $f \in L^1(\mu)$. Alors, pour presque tout point $x$, on a*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x) = \begin{cases} \frac{1}{\mu(X)} \int f \, d\mu & si\ \mu(x) < +\infty, \\ 0 & si\ \mu(x) + \infty; \end{cases}$$

*Preuve.* Le théorème ergodique de Birkhoff nous dit qu'il existe une fonction $f^* \in L^1(\mu)$ invariante telle que $f^*(x) = \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x)$ pour presque tout point $x$. Cette fonction $f^*$ est donc constante presque partout. Si $\mu(X) = +\infty$, elle est nulle puisqu'elle est intégrable ; si $\mu(X) < +\infty$ elle est égale à la moyenne spatiale de $f$ puisque $\int f^* \, d\mu = \int f \, d\mu$. □

**Remarque** On a un résultat similaire pour la convergence quadratique. Si $f \in L^2(\mu)$, la suite $\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i$ converge dans $L^2(\mu)$ vers une fonction invariante. Celle-ci est donc constante, elle est nulle si $\mu(X) = +\infty$, elle est égale à la moyenne spatiale de $f$ si $\mu(X) < +\infty$.

Énonçons encore une caractérisation de l'ergodicité :

Proposition 3.3.4 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré, tel que $\mu(X) = 1$. Les deux propriétés suivantes sont équivalentes :*

**i)** *le système est ergodique;*

**ii)** *pour tout $A$ et $B$ dans $\mathcal{B}$, on a*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(A \cap T^{-i}(B)) = \mu(A)\mu(B).$$

**iii)** *pour tous $f$, $g$ dans $L^2(\mu)$, on a*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int f(g \circ T^i) \, d\mu = \left( \int f \, d\mu \right) \left( \int g \, d\mu \right).$$

*Preuve.* Commençons par prouver que **ii)** implique **i)**. Si **ii)** est vraie, alors pour tout ensemble invariant $A \in \mathcal{B}$, on a

$$\mu(A)^2 = \lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mu(A \cap T^{-i}(A)) = \lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mu(A) = \mu(A),$$

ce qui implique que $\mu(A) = 0$ ou $\mu(A) = 1$. Le système est donc ergodique.

On déduit **ii)** de **iii)** en appliquant cette dernière propriété aux fonctions caractéristiques $\chi_A$ et $\chi_B$. Il reste à prouver que **i)** implique **iii)**. Supposons donc que **i)** soit vraie. Fixons $f$ et $g$ dans $L^2(\mu)$. D'après la remarque précédente, on sait que la suite $\frac{1}{n}\sum_{i=0}^{n-1} g \circ T^i$ converge dans $L^2(\mu)$ vers la fonction constante $\int g\, d\mu$. On en déduit que

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{i=0}^{n-1} \int f(g\circ T^i)\, d\mu = \lim_{n\to+\infty} \int f\left(\frac{1}{n}\sum_{i=0}^{n-1} g\circ T^i\right) d\mu = \int f(\int g\, d\mu)\, d\mu = \int f\, d\mu \int g\, d\mu.$$

□

Introduisons maintenant une notion plus forte.

Proposition 3.3.5 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. Les deux conditions suivantes sont équivalentes :*

**i)** *pour tous $A$, $B$ dans $\mathcal{M}$, on a*

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

**ii)** *pour tous $f$, $g$ dans $L^2(\mu)$, on a*

$$\lim_{n\to+\infty} \int f(g\circ T^n)\, d\mu = \left(\int f\, d\mu\right)\left(\int g\, d\mu\right).$$

*Si ces propriétés sont vérifiées, on dit que le système est* mélangeant.

*Preuve.* On déduit **i)** de **ii)** en appliquant **ii)** aux applications $\chi_A$ et $\chi_B$. On déduit **ii)** de **i)** en remarquant que **ii)** est vraie pour les fonctions étagées, en utilisant ensuite la densité de ces fonctions dans $L^2(\mu)$ ainsi que la continuité de $U_T\ : L^2(\mu) \to L^2(\mu)$. □

Pour un système mélangeant, les évènements "être dans $A$ au temps 0" et "être dans $B$ au temps $n$" sont asymptotiquement indépendants quand $n \to +\infty$. Cette notion est évidemment plus forte que l'ergodicité, puisque pour toute partie invariante $A \in \mathcal{B}$, on a

$$\mu(A)^2 = \lim_{n\to+\infty} \mu(A) \cap T^{-n}(A) = \lim_{n\to+\infty} \mu(A) = \mu(A),$$

(ou si l'on utilise la proposition 3.3.4, puisque la convergence d'une suite implique la convergence au sens de Césaro).

Il existe une propriété intermédiaire entre l'ergodicité et le mélange, que l'on va maintenant introduire :

Proposition 3.3.6 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. Les deux conditions suivantes sont équivalentes :*

**i)** *pour tous $A$, $B$ dans $\mathcal{M}$, on a*

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\left|\mu(A\cap T^{-i}(B))-\mu(A)\mu(B)\right|=0\,;$$

**ii)** *pour tous $f$, $g$ dans $L^2(\mu)$, on a*

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\left|\int f(g\circ T^i)\,d\mu-\left(\int f\,d\mu\right)\left(\int g\,d\mu\right)\right|=0.$$

*Si ces propriétés sont vérifiées, on dit que le système est* faiblement mélangeant*.*

Le fait que les propriétés i) et ii) sont équivalentes se prouve comme dans la proposition 3.3.5 ; le fait que tout système faiblement mélangeant est ergodique se déduit immédiatement de la propriété 3.3.4; le fait que tout système mélangeant est faiblement mélangeant se déduit du théorème de Césaro. Nous allons donner une autre caractérisation utile du mélange faible. Rappelons qu'une partie $I\subset\mathbf{N}$ est de *densité* 0 si

$$\lim_{n\to+\infty}\frac{\sharp(\{0,\ldots,n-1\}\cap I)}{n}=0.$$

Proposition 3.3.7 : *Un système dynamique mesuré $(X,\mathcal{B},\mu,T)$ tel que $\mu(X)=1$ est faiblement mélangeant si et seulement si, pour tous $A$, $B$ dans $\mathcal{M}$, il existe une partie $I\subset\mathbf{N}$ de densité 0 telle que*

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

*Preuve.* La proposition est une conséquence immédiate du lemme suivant.

Lemme 3.3.8 : *Soit $(u_n)_{n\geq 0}$ une suite réelle bornée. Alors les conditions suivantes sont équivalentes :*

**i)** $\displaystyle\lim_{n\to+\infty}\frac{1}{n}\sum_{k=0}^{n-1}|u_k|=0$;

**ii)** *il existe une partie $I\subset\mathbf{N}$ de densité 0 telle que* $\displaystyle\lim_{n\to+\infty,\,n\notin I}u_n=0.$

*Preuve.* Pour toute partie $J\subset\mathbf{N}$ et tout $n\geq 1$, on notera $\alpha_J(n)=\sharp(\{0,\ldots,n-1\}\cap J)$. Commençons par prouver que **i)** implique **ii)**. Pour tout $i\geq 1$, l'ensemble

$$I_i=\{n\geq 0\,|\,|a_n|\geq\frac{1}{i}\}$$

est de densité 0 puisque

$$\frac{1}{n}\sum_{k=0}^{n-1}|u_k|\geq\frac{1}{i}\frac{\alpha_{I_i}(n)}{n}.$$

On peut donc construire dans $\mathbf{N}$ une suite strictement croissante $(n_i)_{i\geq 1}$ telle que pour tout $n\geq n_i$, on ait

$$\frac{\alpha_{I_{i+1}}(n)}{n}\leq\frac{1}{i+1}.$$

Posons
$$I = \bigcup_{i \geq 0} (I_{i+1} \cap [n_i, n_{i+1}[).$$
Il s'agit d'un ensemble de densité zéro car, pour tout $n \in [n_i, n_{i+1}[$, on a
$$I \cap [0, n[ \subset I_{i+1} \cap [0, n[,$$
ce qui implique
$$\frac{\alpha_I(n)}{n} \leq \frac{\alpha_{I_{i+1}}(n)}{n} \leq \frac{1}{i+1}.$$

Remarquons maintenant que si $n \geq n_i$ n'appartient pas à $I$, alors $n$ n'appartient pas à $I_{i+1}$ et donc on a $|a_n| \leq 1/(i+1)$. On a bien montré que $\lim_{n \to +\infty,\, n \notin I} u_n = 0$.

Montrons maintenant que **ii)** implique **i)**. Soit $K > 0$ un majorant de la suite $(|u_n|)_{n \geq 0}$. Pour tout $\varepsilon > 0$, il existe $n_0 \geq 0$ tel que pour tout $n \geq n_0$, on a $\alpha_I(n) \leq \varepsilon n$ et de plus $|u_n| \leq \varepsilon$ si $n \notin I$. On en déduit que pour tout $n \geq \max(n_0, K n_0/\varepsilon)$, on a
$$\begin{aligned}
\frac{1}{n} \sum_{k=0}^{n-1} |u_k| &= \frac{1}{n} \left( \sum_{k<n,\, k \in I} |u_k| + \sum_{k<n_0,\, k \notin I} |u_k| + \sum_{n_0 \leq k < n,\, k \notin I} |u_k| \right) \\
&\leq \frac{K \alpha_I(n)}{n} + K \frac{n_0}{n} + \varepsilon \\
&\leq (K+2)\varepsilon.
\end{aligned}$$

□

Remarquons que l'on déduit également du lemme 3.3.8 le résultat suivant :

Corollaire 3.3.9 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. Les trois conditions suivantes sont équivalentes :*

**i)** *le système est faiblement mélangeant*

**ii)** *pour tous $A$, $B$ dans $\mathcal{M}$, on a*
$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \Big( \mu(A \cap T^{-i}(B)) - \mu(A)\mu(B) \Big)^2 = 0\,;$$

**ii)** *pour tous $f$, $g$ dans $L^2(\mu)$, on a*
$$\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \left| \int f(g \circ T^i)\, d\mu - \left( \int f\, d\mu \right) \left( \int g\, d\mu \right) \right|^2 = 0.$$

## 3.4 Exemples

**Rotations du cercle**.

On note ici $\mathcal{B}$ la tribu borélienne de $\mathbf{T}^1$, $\mu$ la mesure de Haar, et $T_{\widehat{a}}$ la rotation d'angle $\widehat{a} \in \mathbf{T}^1$, où $\widehat{a} = a + \mathbf{Z}$.

Proposition 3.4.1 : *Le système dynamique* $(\mathbf{T}^1, \mathcal{B}, \mu, T_{\widehat{a}})$ *est ergodique si et seulement si* $\widehat{a} \notin \mathbf{Q}/\mathbf{Z}$*. Il n'est jamais mélangeant ni même faiblement mélageant.*

*Preuve.* Considérons $f \in L^2(\mu)$ et développons-la en série de Fourier

$$f(x + \mathbf{Z}) = \sum_{k \in \mathbf{Z}} c_k e^{2i\pi kx}.$$

La série de Fourier de $f \circ T_{\widehat{a}}$ est

$$f \circ T_{\widehat{a}}(x + \mathbf{Z}) = \sum_{k \in \mathbf{Z}} c_k e^{2i\pi ka} e^{2i\pi kx}.$$

L'invariance de $f$ est caractérisée par les égalités :

$$c_k = c_k e^{2i\pi ka},$$

pour $k \in \mathbf{Z}$. Si $a \notin \mathbf{Q}$, on aura $c_k = 0$, si $k \neq 0$, ce qui implique que $f$ est constante. Le système est donc ergodique. Si $a = p/q \in \mathbf{Q}$, l'application $f \ : x \mapsto e^{2i\pi qx}$ est invariante par $T_{\widehat{a}}$ sans être constante, le système n'est pas ergodique.

Considérons $f \ : x + \mathbf{Z} \mapsto e^{2i\pi x}$ et $g \ : x + \mathbf{Z} \mapsto e^{-2i\pi x}$. Remarquons que

$$\left( \int f \, d\mu \right) \left( \int g \, d\mu \right) = 0$$

mais que

$$\lim_{n \to +\infty} \int f(g \circ T^n) \, d\mu \neq 0$$

car

$$\int f(g \circ T^n) \, d\mu = e^{-2i\pi na}.$$

Le système n'est jamais mélangeant. Il n'est pas non plus faiblement mélangeant car

$$\lim \frac{1}{n} \sum_{i=0}^{n-1} \left| \int f(g \circ T^i) \, d\mu \right| = 1 \not\to 0.$$

□

**Rotations du tore $\mathbf{T}^r$.**

Ce qui vient d'être fait se généralise naturellement en dimension plus grande. On note ici $\mathcal{B}$ la tribu borélienne de $\mathbf{T}^r$, $\mu$ la mesure de Haar, et $T_{\widehat{a}}$ la rotation d'angle $\widehat{a} \in \mathbf{T}^1$, où $\widehat{a} = a + \mathbf{Z}^r$ et $a = (a_1, \ldots, a_r) \in \mathbf{R}^r$.

Proposition 3.4.2 : *Le système dynamique* $(\mathbf{T}^1, \mathcal{B}, \mu, T_{\widehat{a}})$ *est ergodique si et seulement si les nombres réels* $1$*,* $a_1$*, …,* $a_r$ *sont rationnellement indépendants. Il n'est jamais mélangeant, ni même faiblement mélangeant.*

*Preuve* Considérons une fonction $f \in L^2(\mu)$ et développons-la en série de Fourier

$$f(x + \mathbf{Z}^r) = \sum_{k \in \mathbf{Z}^r} c_k e^{2i\pi \langle k, x \rangle}.$$

La série de Fourier de $f \circ T_{\widehat{a}}$ est

$$f \circ T_{\widehat{a}}(x + \mathbf{Z}^r) = \sum_{k \in \mathbf{Z}^r} c_k e^{2i\pi\langle k,a\rangle} e^{2i\pi\langle k,x\rangle}.$$

L'invariance de $f$ est cette fois ci caractérisée par les égalités :

$$c_k = c_k e^{2i\pi\langle k,a\rangle}.$$

pour $k \in \mathbf{Z}^r$. Si les nombres 1, $a_1$, ..., $a_r$ sont rationnellement indépendants, on aura $c_k = 0$, si $k \neq 0$, ce qui implique que $f$ est constante. Si les nombres 1, $a_1$, ..., $a_r$ ne sont pas rationnellement indépendants, il existe $k \neq 0$ tel que $\langle k, a\rangle \in \mathbf{Z}$. L'application $f\ : x \mapsto e^{2i\pi\langle k,x\rangle}$ est invariante par $T_{\widehat{a}}$ sans être constante, le système n'est pas ergodique.

Considérons $f\ : x + \mathbf{Z}^r \mapsto e^{2i\pi x_1}$ et $g\ : x + \mathbf{Z}^r \mapsto e^{-2i\pi x_1}$. Là-encore, nous avons

$$\left(\int f\, d\mu\right)\left(\int g\, d\mu\right) = 0$$

mais

$$\lim_{n\to+\infty} \int f(g \circ T^n)\, d\mu \neq 0$$

car

$$\int f(g \circ T^n)\, d\mu = e^{-2i\pi n a_1}.$$

Le système n'est jamais mélangeant. Il n'est pas non plus faiblement mélangeant car

$$\lim \frac{1}{n} \sum_{i=0}^{n-1} \left| \int f(g \circ T^i)\, d\mu \right| = 1 \not\to 0.$$

□

Corollaire 3.4.3 : *Si les nombres* 1, $a_1$,..., $a_r$ *sont rationnellement indépendants, alors* $T_{\widehat{a}}$ *est minimal.*

*Preuve.* Puisque le support de $\mu$ est $\mathbf{T}^r$, le résultat précédent implique que $T_{\widehat{a}}$ est transitif : il existe $\widehat{x}_0$ tel que $\omega(\widehat{x}_0) = \mathbf{T}^r$. Or, on a vu dans la proposition 3.4.1 que la transitivité impliquait la minimalité. □

**Décalage de Bernouilli**.

On se donne un alphabet fini $A$, on munit $A^{\mathbf{N}}$ de la tribu borélienne et on considère le décalage de Bernouilli $\sigma$.

Proposition 3.4.4 : *Soit* $\mu$ *la mesure produit définie par une famille* $(p_a)_{a\in A}$ *de nombres positifs tels que* $\sum_{a\in A} p_a = 1$. *Le système dynamique* $(A^{\mathbf{N}}, \mathcal{B}, \mu, \sigma)$ *est mélangeant.*

*Preuve.* Remarquons que si $C_w^{n_0}$ et $C_{w'}^{n_0'}$ sont deux cylindres, où $w = (w_0, \ldots, w_{m-1})$ et $w' = (w_0', \ldots, w_{m'-1}')$, il existe $N$ tels que les ensembles $\{n_0, \ldots, n_0 + m - 1\}$ et $\{n + n_0', \ldots, n + n_0' + m' - 1\}$ sont disjoints si $n \geq N$. Par construction de $\mu$, on a

$$\mu(C_w^{n_0} \cap \sigma^{-n}(C_{w'}^{n_0'})) = \mu(C_w^{n_0})\mu(C_{w'}^{n_0'}).$$

Pour tout $C_w^{n_0}$, l'ensemble des $Y \in \mathcal{B}$ tels que

$$\lim_{n\to+\infty} \mu(C_w^{n_0} \cap \sigma^{-n}(Y)) = \mu(C_w^{n_0})\mu(Y)$$

est une $\sigma$-algèbre qui contient les cylindres, c'est donc $\mathcal{B}$ tout entier. Ainsi, pour tout $Y \in \mathcal{B}$, l'ensemble des $Z \in \mathcal{B}$ tels que

$$\lim_{n\to+\infty} \mu(Z \cap \sigma^{-n}(Y)) = \mu(Z)\mu(Y)$$

contient tous les cylindres. Puisque c'est une $\sigma$-algèbre, c'est également $\mathcal{B}$. □

PROPOSITION 3.4.5 : *Soit $\mu$ la mesure de Markov définie par une matrice stochastique $M$ et un vecteur fixe $v$. S'il existe une puissance de $M$ dont les coefficients sont strictement positifs, la mesure $\mu$ est mélangeante.*

*Preuve.* Il suffit de prouver que

$$\lim_{n\to+\infty} \mu(C_w^{n_0} \cap \sigma^{-n}(C_{w'}^{n_0'})) = \mu(C_w^{n_0})\mu(C_{w'}^{n_0'})$$

pour tous cylindres $C_w^{n_0}$ et $C_{w'}^{n_0'}$. Supposons que $w = (w_0, \ldots, w_{m-1})$ et $w' = (w_0', \ldots, w_{m'-1}')$. Si $n > m - 1 + n_0 - n_0'$, alors

$$\mu(C_w^{n_0} \cap \sigma^{-n}(C_{w'}^{n_0'})) = v_{w_0} \left(\prod_{k=0}^{m-2} M_{w_k,w_{k+1}}\right) M_{w_{m-1},w_0'}^{n-m+1-n_0+n_0'} \left(\prod_{k=0}^{m'-1} M_{w_k',w_{k+1}'}\right),$$

où on écrit $M^n = (M_{i,j}^n)_{i,j}$, et donc

$$\lim_{n\to+\infty} \mu(C_w^{n_0} \cap \sigma^{-n}(C_{w'}^{n_0'})) = v_{w_0} \left(\prod_{k=0}^{m-2} M_{w_k,w_{k+1}}\right) v_{w_0'} \left(\prod_{k=0}^{m'-1} M_{w_k',w_{k+1}'}\right) = \mu(C_w^{n_0})\mu(C_{w'}^{n_0'}),$$

car on a vu précédemment que $\lim_{n\to+\infty} M_{i,j}^n = v_j$ □

**Remarque** L'équivalent des propositions 5.4.4 et 5.4.5 est vrai pour le décalage bilatéral.

### Endomorphismes linéaires de $\mathbf{T}^1$.

Remarquons la conséquence suivante de la proposition 3.4.4. :

PROPOSITION 3.4.6 : *Le système dynamique $(\mathbf{T}^1, \mathcal{B}, \mu, T)$, où $\mu$ est la mesure de Haar, et $T : \widehat{x} \mapsto p\widehat{x}$, $p \geq 2$, est mélangeant.*

*Preuve.* Nous avons vu qu'il était conjugué au décalage de Bernoulli sur $\{0, \ldots, p-1\}^{\mathbf{N}}$ muni de la mesure équidistribuée. □

### Endomorphismes linéaires de $\mathbf{T}^r$.

Nous allons généraliser le dernier résultat en dimension plus grande. On se donne une application linéaire $A : \mathbf{R}^r \to \mathbf{R}^r$ dont la matrice dans la base canonique de $\mathbf{R}^r$ est à coefficients entiers et on note $\widehat{A}\ : \mathbf{T}^r \to \mathbf{T}^r$ l'endomorphisme de $\mathbf{T}^r$ associé. Rappelons que si $\det A \neq 0$, alors $\widehat{T}$ préserve la mesure de Haar.

PROPOSITION 3.4.7 : *Le système dynamique $(\mathbf{T}^r, \mathcal{B}, \mu, \widehat{A})$ est ergodique si et seulement si les valeurs propres de $A$ ne sont pas racines de l'unité. Dans ce cas, il est mélangeant.*

*Preuve.* Considérons une fonction $f \in L^2(\mu)$ et développons-la en série de Fourier

$$f(x + \mathbf{Z}^r) = \sum_{k \in \mathbf{Z}^r} c_k e^{2i\pi\langle k,x\rangle}.$$

La série de Fourier de $f \circ \widehat{A}$ est

$$f \circ \widehat{A}(x + \mathbf{Z}^r) = \sum_{k \in \mathbf{Z}^r} c_k e^{2i\pi\langle k,Ax\rangle} = \sum_{k \in \mathbf{Z}^r} c_k e^{2i\pi\langle A^t k,x\rangle}.$$

L'invariance de $f$ est donc caractérisée par les égalités :

$$c_k = c_{A^t k},$$

pour $k \in \mathbf{Z}^r$. Dans le cas où les valeurs propres de $A$ ne sont pas racines de l'unité, il en est de même des valeurs propres de $A^t$ et donc, si $k \neq 0$, les indices $(A^t)^n k$, sont tous distincts. Ainsi, si $f$ est invariant et si $k$ est non nul, on aura $c_k = c_{(A^t)^n k}$ et $\sum_{n\geq 1} |c_{(A^t)^n k}|^2 < +\infty$. Ceci n'est possible que si $c_k = 0$. La fonction $f$ sera donc constante.

Dans le cas où l'une au moins des valeurs propres de $A$ est racine de l'unité, il existe $n \geq 1$ tel que $(A^t)^n - \mathrm{Id}_{\mathbf{R}^r}$ n'est pas inversible. Ceci implique qu'il existe $k \in \mathbf{Z}^r$ non nul tel que $(A^t)^n k = k$. La fonction

$$f\ : x + \mathbf{Z}^r \mapsto \sum_{i=0}^{n-1} e^{2i\pi\langle k,A^i x\rangle}$$

est invariante par $\widehat{A}$ sans être constante, le système n'est pas ergodique

Revenons au cas où les valeurs propres de $A$ ne sont pas racines de l'unité. Fixons $k$ et $k'$ dans $\mathbf{Z}^r \setminus \{0\}$ et considérons

$$f\ : x + \mathbf{Z}^r \mapsto e^{\langle k,x\rangle},\ \ g\ : x + \mathbf{Z}^r \mapsto e^{\langle k',x\rangle}.$$

Remarquons que

$$f(x + \mathbf{Z}^r)(g \circ \widehat{A}^n(x + \mathbf{Z}^r)) = e^{\langle k+(A^t)^n k',x\rangle}.$$

Il existe au plus une valeur de $n$ telle que $k + (A^t)^n k' \neq 0$. Ceci implique que

$$\lim_{n\to+\infty} \int f(g \circ T^n)\, d\mu = 0 = \left(\int f\, d\mu\right)\left(\int g\, d\mu\right).$$

L'égalité

$$\lim_{n\to+\infty} \int f(g \circ T^n)\, d\mu = \left(\int f\, d\mu\right)\left(\int g\, d\mu\right)$$

étant évidemment vraie si $f$ ou $g$ est constante, elle est vraie pour tout couple de polynômes trigonométriques. Par densité, on en déduit qu'elle est vraie pour tout couple de fonctions dans $L^2(\mu)$. Le système est bien mélangeant $\square$

**Transformation de Gauss**.

Proposition 3.4.8 : *La transformation de Gauss $x \mapsto \frac{1}{x} - [\frac{1}{x}]$ est ergodique pour la mesure de Gauss.*

*Preuve.* Commençons par prouver le lemme suivant, en gardant les notations introduites en section 5.2, et en notant $\mu$ la mesure de Gauss et $\lambda$ la mesure de Lebesgue.

Lemme 3.4.9 : *Il existe des constantes $C > 0$ et $C' > 0$ telles que, pour tout borélien $A$ de $[0, 1]$ et tout $n$-uplet $(a_1, \ldots, a_n)$ d'entiers strictement positifs, on a*

$$C\mu(I_{a_1,\ldots,a_n})\mu(A) \leq \mu(I_{a_1,\ldots,a_n} \cap T^{-n}(A)) \leq C'\mu(I_{a_1,\ldots,a_n})\mu(A).$$

*Preuve.* Puisque, pour tout borélien $B$, on a

$$\frac{1}{2\ln 2}\lambda(B) \leq \mu(B) \leq \frac{1}{\ln 2}\lambda(B),$$

il suffit de montrer le lemme 5.4.10 en remplaçant $\mu$ par $\lambda$. On a alors

$$\lambda(I_{a_1,\ldots,a_n} \cap T^{-n}(A)) = \int_{\varphi(A)} dx = \int_A |\varphi'(t)|dt,$$

où

$$\varphi \ : t \mapsto \frac{p_n + p_{n-1}t}{q_n + q_{n-1}t},$$

est la réciproque de la restriction de $T^n|_{I_{a_1,\ldots,a_n}}$. Or on a

$$|\varphi'(t)| = \frac{1}{(q_n + q_{n-1}t)^2}.$$

Des inégalités

$$\lambda(I_{a_1,\ldots,a_n}) = \frac{1}{q_n(q_n + q_{n-1})}, \ \ q_{n-1} \leq q_n,$$

on en déduit

$$\frac{1}{2}\lambda(I_{a_1,\ldots,a_n}) \leq |\varphi'(t)| \leq 2\lambda(I_{a_1,\ldots,a_n}),$$

ce qui implique le résultat. □

Démontrons maintenant la proposition 5.4.9. Soit $A$ un borélien vérifiant $T^{-1}(A) = A$. La mesure borélienne $X \mapsto \mu(A \cap X) - C\mu(A)\mu(X)$ prend des valeurs positives sur tout intervalle $I_{a_1,\ldots,a_n}$. Or, la sigma-algèbre engendrée par les $I_{a_1,\ldots,a_n}$ contient les boréliens, la mesure précédente prend donc une valeur positive sur $[0, 1] \setminus A$, ce qui est évidemment impossible. □

En fait l'inégalité à droite donnée dans le lemme pourrait nous permettre de montrer que la transformation de Gauss est mélangeante.

## 3.5 Un peu de théorie spectrale.

Nous nous intéressons dans cette section à des systèmes dynamiques mesurés $(X, \mathcal{B}, \mu, T)$, où $\mu(X) = 1$ et où $T$ est inversible. L'opérateur

$$\begin{aligned} U_T \ : L^2(\mu) &\to L^2(\mu) \\ f &\mapsto f \circ T \end{aligned}$$

est alors un opérateur unitaire : on a $U^* = U^{-1}$ puisque

$$\langle U_T(f), g\rangle = \langle U_T(f), U_T(U_T^{-1}(g))\rangle = \langle f, U_T^{-1}(g)\rangle.$$

Le théorème de Von Neumann nous dit que pour tout $f \in L^2(\mu)$, la suite $\left(\frac{1}{n}\sum_{k=0}^{n-1} U_T^k(f)\right)_{n\geq 0}$ converge dans $L^2(\mu)$ vers $p_{\text{inv}}(f)$, où $p_{\text{inv}}$ est la projection orthogonale sur $L^2_{\text{inv}}(\mu) = \text{Ker}(\text{Id}-U_T)$. L'espace propre $\text{Ker}(\text{Id}-U_T)$ contient toujours les fonctions constantes. L'ergodicité du système signifie que cet espace est de dimension un, il se réduit aux fonctions constantes. Dans ce cas, on a $p_{\text{inv}}(f) = (\int f\, d\mu)\mathbf{1}$. Le caractère mélangeant du système se traduit par le fait que pour tout $f \in L^2(\mu)$ et tout $g \in L^2(\mu)$, on a

$$\lim_{n\to+\infty} \langle U_T^n(f), g\rangle = \lim_{n\to+\infty} \int (f\circ T^n)\,\overline{g}\, d\mu = \left(\int f\, d\mu\right)\left(\int \overline{g}\, d\mu\right) = \langle p_{\text{inv}}(f), g\rangle,$$

c'est-à-dire par la convergence faible de la suite $(U_T^n(f))_{n\geq 0}$ vers $p_{\text{inv}}(f)$. Le théorème qui suit caractérise le mélange faible.

Proposition 3.5.1 : *Le système dynamique $(X, \mathcal{B}, \mu, T)$ est faiblement mélangeant si et seulement si* $1$ *est la seule valeur propre de $U_T$ et* $\dim(\text{Ker}(\text{Id} - U_T)) = 1$.

*Preuve.* Si le système est faiblement mélangeant, il est ergodique et on sait donc que $\dim(\text{Ker}(\text{Id}-U_T)) = 1$. Montrons qu'il n'y a pas d'autre valeur propre que 1. Soit $\lambda \in \mathbf{C}\setminus\{1\}$ et $f \in L^2(\mu)$ tel que $U_T(f) = \lambda f$. Puisque

$$\langle f, \mathbf{1}\rangle = \langle U_T(f), U_T(\mathbf{1})\rangle = \lambda\langle f, \mathbf{1}\rangle,$$

on sait que $\lambda\langle f, \mathbf{1}\rangle = 0$ ; puisque

$$\langle f, f\rangle = \langle U_T(f), U_T(f)\rangle = \lambda\overline{\lambda}\langle f, f\rangle,$$

on sait que $f = 0$ si $|\lambda| \neq 1$. Supposons donc que $|\lambda| = 1$. Le fait que $T$ est faiblement mélangeant implique que

$$0 = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \left|\langle U_T(f), f\rangle - |\langle f, \mathbf{1}\rangle|^2\right| = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} |\langle U_T(f), f\rangle| = \langle f, f\rangle,$$

et donc que $f = 0$.

La proposition réciproque sera basée sur le théorème spectral pour les opérateurs unitaires.

Théorème 3.5.2 : *Soit $U$ un opérateur unitaire sur un espace de Hilbert $\mathcal{H}$. Pour tout $x \in \mathcal{H}$, il existe une unique mesure borélienne de probabilité $\nu_x$ sur le cercle $S^1 = \{z \in \mathbf{C} \mid |z| = 1\}$ telle que pour tout $k \in \mathbf{Z}$, on a*

$$\langle U^k(x), x\rangle = \int z^k\, d\nu_x(z).$$

*L'adhérence $V$ du sous-espace vectoriel engendré par la famille $(U^k(x))_{k\in\mathbf{Z}}$ est alors invariante par $H$ et $H|_V$ est conjugué à l'opérateur $\psi$ défini sur $L^2(\nu_x)$ qui à $g \in L^2(\nu_x)$ associe l'application $z \mapsto zg(z)$.*

*Fin de la preuve de la proposition 3.5.1.* Supposons que 1 soit la seule valeur propre de $U_T$ et que l'espace propre associé soit de dimension 1. Il faut montrer que pour tout $f \in L^2(\mu)$ et $g \in L^2(\mu)$, on a

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \langle U_T(f), g\rangle - \langle f, \mathbf{1}\rangle \overline{\langle g, \mathbf{1}\rangle} \right| = 0.$$

D'après le lemme 3.3.8 il faut montrer que

$$\lim_{n\to\infty} \langle U_T(f), g\rangle = \langle f, \mathbf{1}\rangle \overline{\langle g, \mathbf{1}\rangle}$$

pour $n$ variant dans le complémentaire d'un ensemble de densité 0. Puisque ceci est évidemment vrai si $f$ ou $g$ est constante, il suffit par linéarité de montrer cette égalité si $f$ et $g$ sont dans l'orthogonal $H$ de $\mathrm{Ker}(\mathrm{Id} - U_T)$, égalité qui devient

$$\lim_{n\to\infty} \langle U_T(f), g\rangle = 0.$$

Par linéarité, il suffit de démonter que pour tout $f \in H$, on a

$$\lim_{n\to\infty} \langle U_T(f), f\rangle = 0,$$

pour $n$ variant dans le complémentaire d'un ensemble de densité 0. D'après le lemme 3.3.8, il suffit donc de monter que pour tout $f \in H$, on a

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} |\langle U_T(f), f\rangle|^2 = 0.$$

Notons $\nu_f$ la mesure spectrale associée à $f$ par le théorème 3.5.2. Elle n'a pas d'atome. En effet, si $\nu_f(\{z_0\}) \neq 0$, l'opérateur $\psi$ qui à $g \in L^2(\nu_f)$ associe l'application $z \mapsto zg(z)$ a $z_0$ comme valeur propre puisque $\psi(\chi_{\{z_0\}}) = z_0 \chi_{\{z_0\}}$ et $\chi_{\{z_0\}} \neq 0$. Il en est donc de même de la restriction de $U_T$ à l'adhérence $V$ de l'espace engendré par les itérés $(U_T)^k(f)$, $k \in \mathbf{Z}$. Ceci est impossible puisque cet espace est dans l'orthogonal de $\mathbf{1}$. On veut donc prouver que

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left| \int z^k \, d\nu_f(z) \right|^2 = 0.$$

On écrit

$$\begin{aligned}
\frac{1}{n} \sum_{k=0}^{n-1} \left| \int z^k \, d\nu_f(z) \right|^2 &= \frac{1}{n} \sum_{k=0}^{n-1} \left( \int z^k \, d\nu_f(z) \right) \left( \int \overline{z'}^k \, d\nu_f(z') \right) \\
&= \frac{1}{n} \sum_{k=0}^{n-1} \int \int (z\overline{z'})^k \, d(\nu_f \times \nu_f)(z, z') \qquad . \\
&= \int \int \frac{1}{n} \left( \sum_{k=0}^{n-1} (z\overline{z'})^k \right) \, d(\nu_f \times \nu_f)(z, z')
\end{aligned}$$

Or, on sait que si $z \neq z'$, on a

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} (z\overline{z'})^k = \lim_{n\to+\infty} \frac{1}{n} \left( \frac{1 - (z\overline{z'})^n}{1 - z\overline{z'}} \right) = 0.$$

Puisque $\nu_f$ n'a pas d'atome on en déduit que

$$\lim_{n\to+\infty} \frac{1}{n}\left(\sum_{k=0}^{n-1}(z\overline{z'})^k\right) = 0$$

presque partout (pour $\nu_f \times \nu_f$). Cette suite étant bornée en module par 1, on peut appliquer le théorème de convergence dominée de Lebesgue. □

Le théorème 3.5.2 nous donne une caractérisation spectrale du mélange faible, le système $U_T$ restreint à l'orthogonal des constantes n'a pas de valeur propre : on dit que les système est à *spectre continu.* Le cas opposé est le cas où il existe une base hilbertienne de $L^2(\mu)$ formée de vecteurs propres, on dit alors que le système est à *spectre purement ponctuel.* C'est le cas par exemple des rotations sur $\mathbf{T}^1$, où même sur $\mathbf{T}^r$ (avec la mesure de Haar) puisque la famille $(e_k)_{k\in\mathbf{Z}^r}$ est formée de vecteurs propres, où

$$e_k((x_1,\ldots,x_r)+\mathbf{Z}^r) = e^{2\pi(k_1x_1+\ldots k_rx_r)},$$

si $k=(k_1,\ldots,k_r)\in\mathbf{Z}^r$.

## 3.6 Mesures invariantes des systèmes dynamiques topologiques.

Une application continue $T\ : X\to X$ sur un espace métrique $X$ n'admet pas nécessairement de mesure borélienne invariante (par exemple la translation $n\mapsto n+1$ sur $\mathbf{N}$), contrairement à un homéomorphisme, mais même dans ce cas la mesure n'est pas nécessairement finie (pensons à la translation $k\mapsto k+1$ sur $\mathbf{Z}$). Nous verrons cependant que dans le cas où $X$ est compact, toute application continue $T\ : X\to X$ admet au moins une mesure borélienne de probabilité invariante. Rappelons que le théorème de représentation de Riesz nous donne une bijection entre l'ensemble des mesures boréliennes de probabilité et l'ensemble des formes positives $L$ sur $C(X,\mathbf{C})$ telles que $L(\mathbf{1})=1$ (où $\mathbf{1}$ est la fonction constante égale à 1). Ici $C(X,\mathbf{C})$ représente l'espace vectoriel des applications continues $f\ : X\to\mathbf{C}$ muni de la norme $\|f\|=\sup_{x\in X}|f(x)|$ et $C(X,\mathbf{C})^*$ l'espace des formes linéaires continues $L\ : C(X,\mathbf{C})\to\mathbf{C}$. Une forme est alors positive si elle envoie toute fonction à valeurs réelles positives sur un réel positif (elle est alors nécessairement continue). Il existe une *topologie forte* sur $C(X,\mathbf{C})^*$ définie par la norme

$$\|L\| = \sup_{f\in C(X,\mathbf{C}),\|f\|\le 1}|L(f)|.$$

Il existe également une *topologie faible** définie ainsi : une suite $(L_n)_{n\ge 0}$ converge vers $L$ si $\lim_{n\to+\infty} L_n(f)=L(f)$ pour tout $f\in C(X,\mathbf{C})$.Un résultat fondamental sur cette topologie est que la boule unité $\{L\in C(X,\mathbf{C})^*\,|\,\|L\|\le 1\}$ est compacte (voir exercice ** ). Remarquons également que l'espace $M(X)$ des mesures boréliennes de probabilité est une partie convexe de $C(X,\mathbf{C})^*$, fermée pour la topologie faible* (via la représentation de Riesz), et qu'elle est contenue dans la boule unité puisque $|\int f\,d\mu|\le\int|f|\,d\mu\le\|f\|$ pour toute fonction $f\in C(X,C)$. Ainsi, l'ensemble $M(X)$ muni de la topologie faible* est compact.

Commençons par le théorème de Krylov-Bogolubov, qui est un cas particulier du théorème de Markov-Kakutani.

THÉORÈME 3.6.1 : *Toute application $T\ : X\to X$ laisse invariant au moins une mesure borélienne de probabilité. De plus, l'espace $M_T(X)$ des mesures invariantes est une partie convexe fermée de $M(X)$.*

*Preuve.* On veut prouver que l'application

$$\begin{aligned} T_* \ : M(X) &\to M(X) \\ \mu &\mapsto T_*(\mu) \end{aligned}$$

a un point fixe. Cette application est continue pour la topologie faible*. En effet, si $\lim_{n\to+\infty} \mu_n = \mu$, alors pour tout $f \in C(X, \mathbf{C})$, on a

$$\lim_{n\to+\infty} \int f \, dT_*(\mu_n) = \lim_{n\to+\infty} \int f \circ T \, d\mu_n = \int f \circ T \, d\mu = \int f \, dT_*(\mu).$$

Elle est également affine (elle envoie barycentre sur barycentre). Fixons $\mu \in M(X)$ et posons

$$\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} (T_*)^i(\mu).$$

L'espace $M(T)$ étant compact pour la topologie faible*, on peut extraire une suite $(\mu_{n_m})_{m\geq 0}$ qui converge pour cette topologie vers une mesure $\mu'$. Remarquons que

$$T_*(\mu_{n_m}) - \mu_{n_m} = \frac{1}{n_m}(T_*)^{n_m}(\mu) - \mu$$

et donc que

$$\|T_*(\mu_{n_m}) - \mu_{n_m}\| \leq \frac{2}{n_m}.$$

Cette suite converge donc fortement (et donc également faiblement) vers 0 dans $C(X, \mathbf{C})^*$. Mais elle converge faiblement vers $T_*(\mu') - \mu'$. On en déduit que $T_*(\mu') = \mu'$.

L'ensemble des mesures invariantes est fermé car $T_*$ est continue, il est convexe car $T_*$ est affine. □

Le théorème ne nous dit rien, bien sûr, sur ces mesures invariantes (elles peuvent par exemple être associées à des orbites périodiques). Nous allons voir maintenant comment trouver des mesures invariantes ergodiques. Rappelons qu'un point $x$ d'une partie convexe $C$ d'un espace affine est *extrémal* si, pour tous points $x'$, $x''$ de $C$ et tout $t \in ]0, 1[$, l'égalité $x = tx' + (1-t)x''$ implique que $x' = x'' = x$.

Théorème 3.6.2 : *Les points extrémaux de $M_T(X)$ sont les mesures invariantes ergodiques.*

*Preuve.* Supposons d'abord que $\mu \in M_T(X)$ n'est pas ergodique. Il existe une partie invariante $A$ telle que $0 < \mu(A) < 1$. On peut décomposer

$$\mu = \mu(A)\frac{\mu_A}{\mu(A)} + \mu(X \setminus A)\frac{\mu_{X\setminus A}}{\mu(X \setminus A)}$$

comme barycentre de deux mesures de probabilité invariantes.

Supposons maintenant que $\mu \in M_T(X)$ est ergodique et s'écrit $\mu = t\mu' + (1-t)\mu''$, où $\mu'$, $\mu''$ sont dans $M_T(X)$ et $t$ dans $]0, 1[$. Remarquons que, pour tout $A \in \mathcal{B}$, on a

$$\mu(A) = 0 \Rightarrow \mu'(A) = 0,$$

ce qui implique $\mu' << \mu$. Grâce au théorème de Radon-Nikodym, on peut écrire $d\mu' = \varphi \, d\mu$, où $\varphi \in L^1(\mu)$. Le fait que $\mu'$ et $\mu$ sont invariants par $T$ implique alors que $\varphi$ est également invariant.

On en déduit que $\varphi$ est constante $\mu$-presque partout. Ceci implique que $\varphi = \mathbf{1}$ puisque $\mu$ et $\mu'$ sont des mesures de probabilité, et donc que $\mu' = \mu$. En répétant l'argument, on obtient $\mu'' = \mu$.

Au lieu du théorème de Radon-Nikodym, on peut également invoquer le théorème ergodique de Birkhoff. En effet, pour tout $f \in C(X, \mathbf{C})$, on sait que pour $\mu$-presque tout point $x$, la suite $\displaystyle\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ converge vers $\int f\, d\mu$. Ceci est donc vrai également $\mu'$-presque partout, Si on applique le théorème de Birkhoff à $\mu'$, on sait que pour $\mu'$-presque tout point $x$, la suite $\displaystyle\frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x))$ converge vers une fonction $f^* \in L^1(\mu')$ et que $\int f^*\, d\mu' = \int f\, d\mu'$. Puisque $f^*$ est la fonction constante $\int f\, d\mu$, on obtient $\int f\, d\mu = \int f\, d\mu'$ pour tout $f \in C(X, \mathbf{C})$. □

Corollaire 3.6.3 : *Il existe au moins une mesure borélienne de probabilité invariante qui est ergodique.*

*Preuve.* On sait que $M_T(X)$ est l'adhérence des barycentres des point extrémaux et donc qu'il existe des points extremaux (voir exercice **). □

Le théorème de représentation de Choquet nous donne un résultat plus précis. Pour toute mesure $\mu \in M_T(X)$, il existe une mesure borélienne de probabilité $\nu_\mu$ sur l'espace métrisable $M_T(X)$ tel que l'ensemble $E_T(X)$ des mesures extrémales vérifie $\nu_\mu(E_T(X)) = 1$ et tel que pour tout $f \in C(X, \mathbf{C})$ on a

$$\int f\, d\mu = \int_{M_T(X)} \left( \int f\, d\mu' \right) d\nu_\mu(\mu').$$

Nous allons conclure en introduisant une nouvelle propriété portant sur les systèmes dynamiques topologiques.

Proposition 3.6.4 : *Soit $X$ un espace métrique compact et $T : X \to X$ une application continue. Les conditions suivantes sont équivalentes :*

**i)** *l'ensemble $M_T(X)$ est réduit à un point ;*

**ii)** *pour tout $f \in C(X, \mathbf{C})$ et tout $x \in X$, la suite $\displaystyle\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i(x)$ converge et sa limite ne dépend pas de $x$;*

**iii)** *pour tout $f \in C(X, \mathbf{C})$, la suite $\displaystyle\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i$ converge uniformément vers une fonction constante.*

*Si ces propriétés sont vérifiées, on dira que le système est* uniquement ergodique .

*Preuve.* Expliquons d'abord pourquoi **i)** implique **iii)**. Soit $\mu$ l'unique mesure de probabilité invariante. Supposons qu'il existe $\varepsilon > 0$, une suite d'entiers $(n_m)_{m\geq 0}$ et une suite $(x_m)_{m\geq 0}$ de points de $X$ tels que $\lim_{m\to+\infty} n_m = +\infty$ et

$$\left| \frac{1}{n_m} \sum_{i=0}^{n_m-1} f \circ T^i(x_m) - \int f\, d\mu \right| \geq \varepsilon.$$

Notons $\delta_{x_m}$ la mesure de Dirac en $x_m$ et posons $\mu_m = \frac{1}{n_m} \sum_{i=0}^{n_m - 1} T_*^i \delta_{x_m}$. L'ensemble $M(T)$ étant compact pour la topologie faible*, on peut supposer que la suite $(\mu_m)_{m \geq 0}$ converge faiblement vers une mesure de probabilité $\mu'$. Ici encore, on a

$$T_*(\mu_m) - \mu_m = \frac{1}{n_m} \left( T_*^{n_m}(\delta_{x_m}) - \delta_{x_m} \right)$$

et

$$\| T_*(\mu_m) - \mu_m \| \leq \frac{2}{n_m},$$

ce qui implique que $\mu'$ est invariante. Remarquons maintenant que $\mu' \neq \mu$ car

$$\int f \, d\mu' = \lim_{m \to +\infty} \int f \, d\mu_m \neq \int f \, d\mu.$$

Ceci contredit **i)**.

Il reste à prouver que **ii)** implique **i)**, puisque **iii)** implique trivialement **ii)**. Fixons $x \in X$. Pour tout $f \in C(X, \mathbf{C})$, posons

$$L(f) = \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(x).$$

On obtient une forme positive $L$ sur $C(X, \mathbf{C})$ qui envoie $\mathbf{1}$ sur 1. Elle est donc continue et définit une mesure de probabilité $\mu$. C'est l'unique mesure de probabilité invariante. En effet, le théorème ergodique de Birkhoff nous dit que si $\mu' \in M_T(X)$, alors, pour tout $f \in C(X, \mathbf{C})$, on a

$$\int f \, d\mu' = \int f^* \, d\mu' = \int f \, d\mu.$$

□

**Remarques**

1. Si $T$ est uniquement ergodique; la mesure invariante est nécessairement ergodique.

2. La propriété d'unique ergodicité est stable par conjugaison.

**Exemples**

1. Un endomorphisme linéaire surjectif de $\mathbf{T}^r$ n'est pas uniquement ergodique car il fixe 0 et préserve la mesure de Haar.

2. Les décalages ne sont pas uniquement ergodiques car ils ont une infinité d'orbites périodiques.

3. Une rotation de $\mathbf{T}^1$ d'angle rationnel n'est pas uniquement ergodique car elle a une infinité d'orbites périodiques.

4. Une rotation $T_{\widehat{a}} : \mathbf{T}^r \to \mathbf{T}^r$, où $\widehat{a} = a + \mathbf{Z}^r$ et $a = (a_1, \ldots, a_r) \in \mathbf{R}^r$, telle que 1, $a_1$,..., $a_r$ ne sont pas rationnellement indépendants n'est pas uniquement ergodique car elle préserve une mesure qui n'est pas ergodique, à savoir la mesure de Haar.

5. Une rotation de $\mathbf{T}^1$ d'angle irrationnel est uniquement ergodique, plus généralement une rotation $T_{\widehat{a}} : \mathbf{T}^r \to \mathbf{T}^r$, où $\widehat{a} = a + \mathbf{Z}^r$ et $a = (a_1, \ldots, a_r) \in \mathbf{R}^r$, telle que 1, $a_1$,..., $a_r$ sont rationnellement indépendants. En effet, si $\mu \in M(\mathbf{T}^r)$ est invariante par $T_{\widehat{a}}$, alors $\mu$ est invariante

par tout $T_{n\widehat{a}}$, $n \geq 0$. Puisque la suite $(n\widehat{a})_{n\geq 0}$ est dense dans $\mathbf{T}^r$, on en déduit que $\mu$ est invariante par tout $T_{\widehat{b}}$, $b \in \mathbf{T}^r$. On a vu que la mesure de Haar est la seule mesure vérifiant cette propriété.

6. Un homéomorphisme $F \in \mathrm{Homeo}_+(\mathbf{T}^1)$ de nombre de rotation rationnel est uniquement ergodique si et seulement s'il n'a qu'une orbite périodique (voir exercice **)

7. Un homéomorphisme $F \in \mathrm{Homeo}_+(\mathbf{T}^1)$ de nombre de rotation irrationnel $\widehat{\rho}$ est uniquement ergodique. C'est évident dans le cas où $F$ est conjugué à $T_{\widehat{\rho}}$. Étudions l'autre cas, en notant $H$ la semi-conjugaison de $F$ sur $T_{\widehat{\rho}}$. Soit $\mu \in M_F(\mathbf{T}^1)$. On sait que $H_*(\mu)$ est la mesure de Haar $\mu_0$ car $T_{\widehat{\rho}}$ est uniquement ergodique. Le théorème de récurrence de Poincaré nous dit que l'adhérence de toute composante connexe $I$ du complémentaire de $X = \Omega(F)$ est de mesure nulle car elle est errante. On en déduit en particulier que le support de $\mu$ est dans $X$ (et même égal à $X$ car il est invariant). Ces intervalles sont les préimages de $H$ qui ne sont pas réduites à un point et il y en a un nombre dénombrable. La réunion de ces intervalles s'écrit $H^{-1}(Y)$ où $Y$ est l'ensemble dénombrable des points de $\mathbf{T}^1$ ayant plus d'un antécédent. On a $\mu_0(Y) = \mu(H^{-1}(Y)) = 0$. L'application $H$ induit une bijection entre $\mathbf{T}^1 \setminus H^{-1}(Y)$ et $\mathbf{T}^1 \setminus Y$ qui transporte $\mu$ sur $\mu_0$.

Nous allons donner un exemple intéressant de système dynamique uniquement ergodique, puis en déduire des propriétés d'équirépartition de certaines suites

PROPOSITION 3.6.5 : *Soit $\widehat{a} \in \mathbf{Q}/\mathbf{Z}$., alors l'homéomorphisme*

$$\begin{aligned} f \ &: \mathbf{T}^r \to \mathbf{T}^r \\ (\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_r) &\mapsto (\widehat{x}_1 + \widehat{a}, \widehat{x}_1 + \widehat{x}_2, \ldots, \widehat{x}_{r-1} + \widehat{x}_r) \end{aligned}$$

*est uniquement ergodique.*

*Preuve.* L'homéomorphisme $f$, composée d'une translation et d'un automorphisme linéaire de $\mathbf{T}^r$, préserve la mesure de Lebesgue. Nous voulons montrer que c'est la seule mesure invariante. Rappelons que pour tout $k = (k_1, \ldots, k_r) \in \mathbf{Z}^r$, on définit une fonction $e_k \ : \mathbf{T}^r \to \mathbf{C}$ en posant

$$e_k((x_1, \ldots, x_r) + \mathbf{Z}^r) = e^{2\pi(k_1 x_1 + \ldots k_r x_r)}.$$

Donnons nous une mesure de probabilité $\mu$ invariante et, pour tout $k \in \mathbf{Z}^r$, considérons le $k$-ième coefficient de Fourier $c_k(\mu) = \int_{\mathbf{T}^r} e_{-k}\, d\mu$. L'invariance de $\mu$ se traduit par l'égalité

$$c_{(k_1,\ldots,k_r)}(\mu) = \int_{\mathbf{T}^r} e_{-k}\, d\mu = \int_{\mathbf{T}^r} e_{-k} \circ f\, d\mu = e^{-2i\pi k_1 \alpha} c_{(k_1+k_2,\ldots,k_{r-1}+k_r,k_r)}(\mu)$$

Il nous faut bien sûr prouver que $c_k(\mu) = 0$ si $k \neq 0$. L'égalité au-dessus nous dit que c'est le cas si $k = (k_1, 0, \ldots, 0)$. Montrons par récurrence sur $s$ que c'est le cas si $k = (k_1, \ldots, k_s, 0, \ldots, 0)$. Supposons donc que ceci est vrai pour $s-1$. Fixons $k = (k_1, \ldots, k_s, 0, \ldots, 0)$. Remarquons que, pour tout $n \geq 0$, la fonction $e_k \circ f^n$ s'écrit

$$e_k \circ f^n = \lambda_n e_{k(n)},$$

où

$$k(n) = (k_1(n), \ldots, k_{s-1}(n), k_s, 0, \ldots, 0).$$

Remarquons également que si $n' \neq n$, alors $k(n') \neq k(n)$, ce qui implique, grâce à l'hypothèse de récurrence que

$$\int_{\mathbf{T}^r} (e_k \circ f^n)(\overline{e_k \circ f^{n'}})\, d\mu = \int_{\mathbf{T}^r} \lambda_n \overline{\lambda_{n'}} e_{k(n)-k(n')}\, d\mu = \lambda_n \overline{\lambda_{n'}} c_{k(n')-k(n)}(\mu) = 0.$$

Ainsi la famille $(e_k \circ f^n)_{n\geq 0}$ est orthonormée dans $L^2(\mu)$. On en déduit, par l'inégalité de Bessel, que

$$\sum_{n=0}^{\infty} |\langle e_0, e_k \circ f^n \rangle_{L^2(\mu)}|^2 \leq \|e_0\|^2_{L^2(\mu)} = 1.$$

Or, on a

$$\langle e_0, e_k \circ f^n \rangle_{L^2(\mu)} = \int_{\mathbf{T}^r} \overline{e_k \circ f^n}\, d\mu = \int_{\mathbf{T}^r} \overline{e_k}\, d\mu = c_k(\mu),$$

ce qui implique que $c_k(\mu) = 0$. □

Nous dirons qu'une suite de réels $(x_n)_{n\geq 0}$ est *équirépartie modulo un*, si pour tout intervalle $[a,b]$ de $[0,1]$, on a

$$\lim_{n\to+\infty} \frac{1}{n} \sharp\{k < n \,|\, x_k \in [a,b]\} = b - a.$$

Si on pose $\widehat{x}_n = x_n$, ceci signifie que la suite $(\mu_n)_{n\geq 1}$ converge vers la mesure de Haar pour la topologie faible*, où $\mu_n = \dfrac{1}{n} \displaystyle\sum_{k<n} \delta_{\widehat{x}_k}$. De façon équivalente, une suite de réels $(x_n)_{n\geq 0}$ est équirépartie modulo un, si pour toute fonction continue $\varphi\ : \mathbf{R} \to \mathbf{R}$, périodique de période 1, on a

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(x_k) = \int_0^1 f(x)\, dx.$$

Pour qu'il en soit ainsi, il suffit qu'elle vérifie le *critère de Weyl*, c'est-à-dire il suffit que pour tout entier $m \neq 0$, on ait

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} e^{2i\pi m x_k} = 0.$$

Ainsi par exemple, on sait que pour tout nombre irrationnel $\alpha$, la suite $(n\alpha)_{n\geq 0}$ est équirépartie. On vérifie aisément que cette suite vérifie le critère de Weyl, mais on peut également invoquer l'unique ergodicité de la rotation $\widehat{x} \mapsto \widehat{x}+\widehat{\alpha}$, où $\widehat{\alpha} = a+\mathbf{Z}$. La proposition 3.5.5 va nous permettre de généraliser cette situation et construire d'autres suites équiréparties plus complexes.

Corollaire 3.6.6 : *Soit $P$ un polynôme réel non constant à coefficient directeur irrationnel. Alors la suite $P(n)_{n\geq 0}$ est équirépartie modulo 1*

*Preuve.* On écrit $P(X) = a_0 + a_1 X + \ldots + a_r X^r$, où $a_r$ est irrationnel. On définit alors une suite $(P_i)_{0\leq i\leq r}$ de polynômes par la relation de récurrence $P_i(X) = P_{i+1}(X+1) - P_{i+1}(X)$ et la condition $P_r = P$. Le polynôme $P_i$ est de degré $i$ et de coefficient directeur $r(r-1), \ldots (i+1)a_r$. Posons $\widehat{x}^n = (P_1(n), P_2(n), \ldots, P_r(n)) + \mathbf{Z}^r$ et remarquons que $\widehat{x}^{n+1} = f(\widehat{x}^n)$, où

$$\begin{aligned} f\ &: \mathbf{T}^r \to \mathbf{T}^r \\ (\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_r) &\mapsto (\widehat{x}_1 + \widehat{a}, \widehat{x}_1 + \widehat{x}_2, \ldots, \widehat{x}_{r-1} + \widehat{x}_r) \end{aligned}$$

et $\widehat{a} = r! a_r$. Pour toute fonction continue $\varphi\ : \mathbf{T}^1 \to \mathbf{R}$, on peut définir $\Phi\ : (\widehat{x}_1, \ldots, \widehat{x}_r) \mapsto \varphi(\widehat{x}_r)$.

Puisque $f$ est uniquement ergodique, on obtient

$$\begin{aligned}
\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(P(k)+\mathbf{Z}) &= \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \Phi(\widehat{x}^k) \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \Phi(f^k(\widehat{x}0)) \\
&= \int_{\mathbf{T}^r} \Phi(\widehat{x}_1,\ldots,\widehat{x}_r)\, d\widehat{x}_1 \ldots d\widehat{x}_r \\
&= \int_{\mathbf{T}^1} \varphi(\widehat{x}_r)\, dx_r
\end{aligned}$$

□

## 4.1 Introduction

Les décalages de Bernouilli $\sigma_p$ et $\sigma_q$, définis respectivement sur $\{1,\ldots,p\}^{\mathbf{N}}$ et $\{1,\ldots,q\}^{\mathbf{N}}$ ne sont pas topologiquement conjugués si $p \neq q$, il suffit de remarquer par exemple qu'il n'ont pas le même nombre de points fixes. Si on munit chacun des ensembles de la mesure équidistribuée, on obtient deux systèmes dynamiques mesurés et on peut se demander si ces deux systèmes sont conjugués (les points périodiques formant de chaque côté un ensemble dénombrable et donc de mesure nulle peuvent être oubliés). En 1958, Kolmogorov a prouvé que ce n'était pas le cas, en utilisant la notion d'*entropie*, inspirée des travaux de Shanon sur la théorie de l'information, notion améliorée ensuite par Sinai en 1959. On va s'intéresser ici aux principales propriétés de l'entropie.

## 4.2 Entropie d'une partition

On suppose dans cette section que l'ensemble $X$ est muni d'une $\sigma$-algèbre $\mathcal{B}$ et d'une mesure de probabilité $\mu$ sur $\mathcal{B}$. Une *partition mesurable* de $X$ est une famille finie $\mathcal{P} = (P_i)_{i\in I}$ d'éléments de $\mathcal{B}$ tels que

- $X = \bigcup_{i\in I} P_i$ ;
- $\mu(P_i \cap P_{i'}) = 0$ si $i \neq i'$..

Pour tout ensemble $P \in \mathcal{B}$ on écrira $P \in \mathcal{P}$, s'il existe $i \in I$ tel que $P = P_i$.

On dira qu'une partition $\mathcal{Q}$ est *plus fine* que $\mathcal{P}$ ou (*raffine* $\mathcal{P}$) et on écrira $\mathcal{P} \preceq \mathcal{Q}$, si pour tout $Q \in \mathcal{Q}$ il existe $P \in \mathcal{P}$ tel que $\mu(Q \setminus P) = 0$. On obtient un préordre sur l'ensemble des partitions mesurables et on notera $\sim$ la relation d'équivalence associée. Remarquons que deux partitions $\mathcal{P}$ et $\mathcal{Q}$ sont équivalentes si et seulement si, pour tout $Q \in \mathcal{Q}$ il existe $P \in \mathcal{P}$ tel que $\mu(Q\Delta P) = 0$. Remarquons également que toute partition mesurable $Q$ est équivalente à une partition mesurable $\mathcal{P} = (P_i)_{i\in I}$, où les $P_i$ sont disjoints deux à deux (on aurait pu, dans ce qui suit, se restreindre à ces partitions). Remarquons également qu'il existe une bijection naturelle entre ces partitions (au sens fort) et les sous-algèbres finies de $\mathcal{B}$. Enfin, notons que toute partition raffine la partition $\{X\}$ constituée de l'unique élément $X$.

Si $(\mathcal{P}^j)_{1\leq j\leq n}$ est une famille finie de partitions mesurables de $X$, on peut définir la partition mesurable $\bigvee_{j=1}^n \mathcal{P}^j$ dont les éléments sont les $\bigcap_{1\leq j\leq n} P^j$, où $P^j \in \mathcal{P}^j$. Elle raffine chaque $\mathcal{P}^j$.

Considérons maintenant la fonction

$$\begin{aligned} \phi \ &: [0,1] \to [0,+\infty[ \\ &x \neq 0 \mapsto -x\ln x \ . \\ &\quad 0 \mapsto 0 \end{aligned}$$

Elle est continue, concave, s'annule en 0 et 1 et atteint son maximum $1/e$ en $1/e$.

**Définition.** L'*entropie* d'une partition $\mathcal{P} = (P_i)_{i\in I}$ est

$$H(\mathcal{P}) = \sum_{i\in I} \phi(\mu(P_i)) = \sum_{i\in I} -\mu(P_i)\ln\mu(P_i).$$

C'est la moyenne de la *fonction information* $I_{\mathcal{P}}$ définie presque partout par l'égalité $I_{\mathcal{P}}(x) = -\ln\mu(P_i)$ si $x \in P_i$. Ceci rend naturelle la définition suivante de l'entropie conditionnelle d'une partition par rapport à une autre partition.

**Définition.** Si $\mathcal{P} = (P_i)_{i\in I}$ et $\mathcal{Q} = (Q_j)_{j\in J}$ sont deux partitions mesurables, l'*entropie de* $\mathcal{P}$ *relativement à* $\mathcal{Q}$ est

$$H(\mathcal{P}|\mathcal{Q}) = \sum_{i\in I, j\in J'} \mu(Q_j)\phi\left(\frac{\mu(P_i\cap Q_j)}{\mu(Q_j)}\right) = \sum_{i\in I, j\in J'} -\mu(P_i\cap Q_j)\ln\frac{\mu(P_i\cap Q_j)}{\mu(Q_j)}$$

où $J' = \{j\in J \,|\, \mu(Q_j)\neq 0\}$.

**Remarque** On peut vérifier que $H(\mathcal{P}|\mathcal{Q}) = 0$ si et seulement si $P \preceq Q$, en particulier on a $H(\mathcal{P}|\mathcal{P}) = 0$. On peut vérifier également que $H(\mathcal{P}, \{X\}) = H(\mathcal{P})$.

Énonçons les principales propriétés de l'entropie.

PROPOSITION 4.2.1 : *Supposons que* $\mathcal{P} = (P_i)_{i\in I}$, $\mathcal{Q} = (Q_j)_{j\in J}$ *et* $\mathcal{S} = (S_k)_{k\in K}$ *sont des partitions mesurables. On a alors les propriétés suivantes :*

**i)** $H(\mathcal{P}) \leq \ln\sharp I$ *avec égalité si et seulement si tous les* $\mu(P_i)$ *sont égaux ;*

**ii)** $H(\mathcal{P}|\mathcal{Q}\vee\mathcal{S}) \leq H(\mathcal{P}|\mathcal{S})$ *et* $H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P})$ *;*

**iii)** $H(\mathcal{P}\vee\mathcal{Q}|\mathcal{S}) = H(\mathcal{P}|\mathcal{Q}\vee\mathcal{S}) + H(\mathcal{Q}|\mathcal{S})$ *et* $H(\mathcal{P}\vee\mathcal{Q}) = H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q})$ *;*

**iv)** *si* $Q \preceq P$, *alors* $H(\mathcal{Q}|\mathcal{S}) \leq H(\mathcal{P}|\mathcal{S})$ *et* $H(\mathcal{Q}) \leq H(\mathcal{P})$ *;*

**v)** *si* $S \preceq Q$, *alors* $H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P}|\mathcal{S})$ *;*

**vi)** $H(\mathcal{P}) \leq H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q})$ *;*

**vii)** $H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P}|\mathcal{S}) + H(\mathcal{S}|\mathcal{Q})$.

*Démonstration.* L'assertion **i)** est une conséquence de la concavité stricte de la fonction ln. Posons $I' = \{i\in I \,|\, \mu(P_i) > 0\}$. On a

$$H(\mathcal{P}) = \sum_{i\in I'} -\mu(P_i)\ln\mu(P_i) = \sum_{i\in I'}\mu(P_i)\ln\frac{1}{\mu(P_i)} \leq \ln\left(\sum_{i\in I'}\mu(P_i)\frac{1}{\mu(P_i)}\right) = \ln(\sharp I') \leq \ln(\sharp I).$$

L'égalité a lieu quand chaque inégalité est une égalité, c'est-à-dire quand on a $\mu(P_i) = \dfrac{1}{\sharp I}$, pour tout $i\in I$.

L'assertion **ii)** est une conséquence de la concavité de $\phi$. Fixons $i \in I$ et $k \in K' = \{k \in K \,|\, \mu(S_k)\neq 0\}$, et posons $J_k = \{j\in J \,|\, \mu(Q_j\cap S_k)\neq 0\}$. On a

$$\begin{aligned}\sum_{j\in J_k}\mu(Q_j\cap S_k)\phi\left(\frac{\mu(P_i\cap Q_j\cap S_k)}{\mu(Q_j\cap S_k)}\right) &\leq \mu(S_k)\phi\left(\sum_{j\in J_k}\frac{\mu(Q_j\cap S_k)}{\mu(S_k)}\frac{\mu(P_i\cap Q_j\cap S_k)}{\mu(Q_j\cap S_k)}\right)\\ &= \mu(S_k)\phi\left(\frac{\mu(P_i\cap S_k)}{\mu(S_k)}\right).\end{aligned}$$

En sommant sur $i$ et $k$ on obtient

$$H(\mathcal{P}|\mathcal{Q}\vee\mathcal{S}) \leq H(\mathcal{P}|\mathcal{S}).$$

Appliquant ce qui précède à $S = \{X\}$, on obtient

$$H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P}).$$

Pour établir **iii)**, écrivons

$$\begin{aligned}
&H(\mathcal{P} \vee \mathcal{Q}|\mathcal{S}) \\
&= \sum_{i,j,k} -\mu(P_i \cap Q_j \cap S_k) \ln\left(\frac{\mu(P_i \cap Q_j \cap S_k)}{\mu(S_k)}\right) \\
&= \sum_{i,j,k} -\mu(P_i \cap Q_j \cap S_k) \ln\left(\frac{\mu(P_i \cap Q_j \cap S_k)}{\mu(Q_j \cap S_k)}\right) + \sum_{i,j,k} -\mu(P_i \cap Q_j \cap S_k) \ln\left(\frac{\mu(Q_j \cap S_k)}{\mu(S_k)}\right) \\
&= H(\mathcal{P}|\mathcal{Q} \vee \mathcal{S}) + H(\mathcal{Q}|\mathcal{S}).
\end{aligned}$$

Si on applique cette dernière égalité à $S = \{X\}$, on trouve

$$H(\mathcal{P} \vee \mathcal{Q}) = H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q}).$$

Pour obtenir **iv)**, il suffit de remarquer que si $\mathcal{Q} \preceq \mathcal{P}$, alors $\mathcal{P} \vee \mathcal{Q} = \mathcal{P}$, et on a donc

$$H(\mathcal{P}|\mathcal{S}) = H(\mathcal{P} \vee \mathcal{Q}|\mathcal{S}) = H(\mathcal{P}|\mathcal{Q} \vee \mathcal{S}) + H(\mathcal{Q}|\mathcal{S}) \geq H(\mathcal{Q}|\mathcal{S}).$$

Dans le cas où $S = \{X\}$, on obtient

$$H(\mathcal{P}) \geq H(\mathcal{Q}).$$

Pour obtenir **v)**, il suffit de remarquer que si $\mathcal{S} \preceq \mathcal{Q}$, alors $\mathcal{Q} \vee \mathcal{S} = \mathcal{Q}$, et on a donc

$$H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P}|\mathcal{Q} \vee \mathcal{S}) \leq H(\mathcal{P}|\mathcal{S}).$$

L'assertion **vi)** se déduit de

$$H(\mathcal{P}) \leq H(\mathcal{P} \vee \mathcal{Q}) = H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}),$$

et l'assertion **vii)** de

$$H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P} \vee \mathcal{S}|\mathcal{Q}) = H(\mathcal{P}|\mathcal{Q} \vee \mathcal{S}) + H(\mathcal{S}|\mathcal{Q}) \leq H(\mathcal{P}|\mathcal{S}) + H(\mathcal{S}|\mathcal{Q}).$$

□

On en déduit immédiatement :

Corollaire 4.2.2 : *On obtient une distance $D$ sur l'ensemble des classes d'équivalence de la relation $\sim$, appelée la* distance de Rokhlin*, en posant :*

$$D(\mathcal{P}, \mathcal{Q}) = H(\mathcal{P}, \mathcal{Q}) + H(\mathcal{Q}, \mathcal{P}),$$

*pour toutes partitions mesurables $\mathcal{P}$ et $\mathcal{Q}$.*

Concluons cette section par un résultat qui sera très utile plus tard. Pour toute suite croissante $(\mathcal{P}_m)_{m\geq 0}$ de partitions mesurables de $X$, notons $\bigvee_{m=0}^{+\infty} \mathcal{P}_m$ la $\sigma$-algèbre engendrée par

$\bigcup_{m\geq 0}\mathcal{P}_m$ et $\overline{\bigvee_{m=0}^{+\infty}\mathcal{P}_m}$ la $\sigma$-algèbre formée des éléments $A\in\mathcal{B}$ tels qu'il existe $B\in\bigvee_{m=0}^{+\infty}\mathcal{P}_m$ vérifiant $\mu(A\Delta B)=0$. On dira que la famille $(\mathcal{P}_m)_{m\geq 0}$ est *génératrice* si $\overline{\bigvee_{m=0}^{+\infty}\mathcal{P}_m}=\mathcal{B}$.

PROPOSITION 4.2.3 : *Si $(\mathcal{P}_m)_{m\geq 0}$ est une suite croissante génératrice de $X$, alors pour toute partition mesurable $\mathcal{P}$ on a $\lim_{m\to+\infty}H(\mathcal{P},\mathcal{P}_m)=0$.*

*Démonstration.* Commençons par quelques lemmes.

LEMME 4.2.4 : *Pour tout $A\in\mathcal{B}$, il existe un entier $m\geq 0$ et un ensemble $B\in\mathcal{B}$, réunion d'éléments de $\mathcal{P}_m$, tels que $\mu(A\Delta B)<\varepsilon$.*

*Démonstration.* Il faut montrer que l'ensemble $\mathcal{C}$ des éléments $A\in\mathcal{B}$ qui vérifie la condition du lemme est une $\sigma$-algèbre qui contient tous les $\mathcal{P}_m$ et qui vérifie la condition suivante :

- si $A\in\mathcal{C}$ et $A'\in\mathcal{B}$ vérifient $\mu(A\Delta A')=0$, alors $A'\in\mathcal{C}$.

Le fait que $\mathcal{C}$ contient chaque $\mathcal{P}_m$ est évident, la fait que la condition précédente est vérifiée se déduit immédiatement de la relation ensembliste suivante :

$$A'\Delta B\subset(A'\Delta A)\cup(A\Delta B).$$

Prouvons maintenant que $\mathcal{C}$ est une $\sigma$-algèbre. Cet ensemble contient $X$ et est stable par passage au complémentaire puisque $(X\setminus A)\Delta(X\setminus B)=A\Delta B$. Il reste à montrer qu'il est stable par union dénombrable. Soit $(A_r)_{r\geq 1}$ une famille d'éléments de $\mathcal{B}$ et $\varepsilon>0$. On peut trouver $r_0$ tel que

$$\mu\left(\bigcup_{r\geq 1}A_r\setminus\bigcup_{1\leq r\leq r_0}A_r\right)\leq\frac{\varepsilon}{2}.$$

Pour tout $r\in\{1,\dots,r_0\}$, on peut trouver un entier $m_r\geq 0$ et un ensemble $B_r$, réunion d'éléments de $\mathcal{P}_{m_r}$, tel que $\mu(A_r\Delta B_r)<\varepsilon/2r_0$. L'ensemble $\bigcup_{1\leq r\leq r_0}B_r$ est une union d'éléments de $\mathcal{P}_m$, où $m=\sup_{1\leq r\leq r_0}m_r$. Remarquons maintenant que

$$\left(\bigcup_{r\geq 1}A_r\right)\Delta\left(\bigcup_{1\leq r\leq r_0}B_r\right)\subset\left(\bigcup_{r\geq 1}A_r\setminus\bigcup_{1\leq r\leq r_0}A_r\right)\cup\left(\bigcup_{1\leq r\leq r_0}A_r\Delta B_r\right),$$

ce qui implique que

$$\mu\left(\left(\bigcup_{r\geq 1}A_r\right)\Delta\left(\bigcup_{1\leq r\leq r_0}B_r\right)\right)\leq\varepsilon.$$

□

LEMME 4.2.5 : *Pour tout $r\geq 1$ et tout $\varepsilon>0$, il existe $\eta>0$ tel que pour toutes partitions mesurables $\mathcal{P}=(P_i)_{i\in I}$, $\mathcal{Q}=(Q_i)_{i\in I}$ vérifiant $\sharp I=r$ et $\mu(P_i\Delta Q_i)\leq\eta$, on a $H(\mathcal{P}|\mathcal{Q})\leq\varepsilon$.*

*Démonstration.* Choisissons $\eta$ assez petit pour que $\phi$ soit croissante sur $[0,\eta]$ et décroissante sur $[1-r\eta,1]$. Écrivons

$$H(\mathcal{P}|\mathcal{Q})=\sum_{i\neq j}\mu(Q_j)\phi\left(\frac{\mu(P_i\cap Q_j)}{\mu(Q_j)}\right)+\sum_i\mu(Q_i)\phi\left(\frac{\mu(P_i\cap Q_i)}{\mu(Q_i)}\right).$$

Si $i \neq j$, alors $\mu(P_i \cap Q_j) \leq \mu(P_i \Delta Q_i) \leq \eta$. Ainsi, on a

$$\begin{aligned}
\mu(Q_j)\phi\left(\frac{\mu(P_i \cap Q_j)}{\mu(Q_j)}\right) &= -\mu(P_i \cap Q_j) \ln\left(\frac{\mu(P_i \cap Q_j)}{\mu(Q_j)}\right) \\
&= -\mu(P_i \cap Q_j) \ln \mu(P_i \cap Q_j) + \mu(P_i \cap Q_j) \ln \mu(Q_j) \\
&\leq -\mu(P_i \cap Q_j) \ln \mu(P_i \cap Q_j) \\
&\leq -\eta \ln(\eta) = \phi(\eta).
\end{aligned}$$

Remarquons maintenant que

$$1 - \sum_i \mu(P_i \cap Q_i) = \sum_i \mu(P_i) - \mu(P_i \cap Q_i) = \sum_i \mu(P_i \setminus P_i \cap Q_i) \leq r\eta.$$

Le fait que $\phi$ est concave implique que

$$\sum_i \mu(Q_i)\phi\left(\frac{\mu(P_i \cap Q_i)}{\mu(Q_i)}\right) \leq \phi\left(\sum_i \mu(P_i \cap Q_i)\right) \leq \phi(1 - r\eta).$$

On a donc

$$H(\mathcal{P}|\mathcal{Q}) \leq r(r-1)\phi(\eta) + \phi(1 - r\eta).$$

Si $\eta$ est suffisamment petit, alors le terme à droite est inférieur ou égal à $\varepsilon$. □

*Démonstration de la proposition 4.2.3.* Soit $\mathcal{P} = (P_i)_{1 \leq i \leq r}$ une partition mesurable. Donnons nous $\varepsilon > 0$. Commençons par considèrer le réel $\eta$ defini par le lemme 4.2.5. Fixons ensuite $\delta$ suffisamment petit pour que l'on ait

$$(1 + r(r-1))\delta \leq \eta$$

et

$$(r-1)(1 + r(r-1))\delta \leq \eta,$$

(la seconde condition implique bien sûr la première si $r > 1$).

On peut trouver un entier $m_0 \geq 0$ et pour tout $i \in \{1 \ldots r\}$, un ensemble $B_i$, union d'éléments de $\mathcal{P}_{m_0}$, tel que $\mu(P_i \Delta B_i) < \delta$. On va construire une partition $\mathcal{Q} = (Q_i)_{1 \leq i \leq r}$, où chaque $Q_i$ est une union d'éléments de $\mathcal{P}_{m_0}$, tel que $\mu(P_i \Delta Q_i) \leq \eta$. Puisqu'on a $\mathcal{Q} \preceq \mathcal{P}_m$, pour tout $m \geq m_0$, on en déduit que $H(\mathcal{P}, \mathcal{P}_m) \leq H(\mathcal{P}, \mathcal{Q}) \leq \varepsilon$.

Remarquons que si $i \neq j$, alors à un ensemble de mesure nulle près on a

$$B_i \cap B_j \subset (P_i \Delta B_i) \cup (P_j \Delta B_j),$$

ce qui implique que

$$\mu(B_i \cap B_j) \leq 2\delta.$$

On en déduit que l'ensemble

$$B = \cup_{i \neq j} B_i \cap B_j$$

vérifie

$$\mu(B) \leq r(r-1)\delta.$$

Considérons la partition $\mathcal{Q} = (Q_i)_{1 \leq i \leq r}$ où

$$\begin{cases} Q_i = B_i \setminus B & \text{si } i < r, \\ Q_r = X \setminus \bigcup_{i<r} B_i & , \end{cases}$$

et montrons que $\mu(P_i \Delta Q_i) \leq \eta$ pour tout $i \in \{1, \dots, r\}$. Si $i < r$, alors

$$P_i \Delta Q_i \subset (P_i \Delta B_i) \cup (B_i \Delta Q_i) \subset (P_i \Delta B_i) \cup B,$$

ce qui implique que

$$\mu(P_i \Delta Q_i) \leq \delta + r(r-1)\delta \leq \eta.$$

Le fait que $\mathcal{P}$ et $\mathcal{Q}$ sont des partitions implique que

$$P_r \Delta Q_r \subset \bigcup_{1 \leq i < r} P_i \Delta Q_i,$$

et donc que

$$\mu(P_r \Delta Q_r) \leq (r-1)(1 + r(r-1)\delta) \leq \eta.$$

□

### 4.3 Entropie d'un système dynamique

On se donne maintenant un système dynamique mesuré $(X, \mathcal{B}, \mu, T)$, où $\mu$ est une mesure de probabilité. Remarquons que pour toute partition mesurable $\mathcal{P} = (P_i)_{i \in I}$, on obtient une autre partition mesurable $T^{-1}(\mathcal{P}) = (T^{-1}(P_i))_{i \in I}$. Remarquons également que $h(T^{-1}(\mathcal{P})) = h(\mathcal{P})$ et plus généralement que $h(T^{-1}(\mathcal{P}) | T^{-1}(\mathcal{Q})) = h(\mathcal{P}|\mathcal{Q})$ si $\mathcal{Q}$ est une autre partition mesurable.

Proposition 4.3.1 : *Si $\mathcal{P}$ est une partition mesurable de $X$, alors la suite*

$$\frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)$$

*est convergente. Sa limite $h(T, \mathcal{P})$ est* l'entropie de $T$ relativement à $\mathcal{P}$, *elle vérifie :*

$$0 \leq h(T, \mathcal{P}) \leq H(\mathcal{P}).$$

*Démonstration.* La suite $(u_n)_{n \geq 1}$, où $u_n = H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)$, est sous-additive. En effet, on a

$$\begin{aligned}
u_{n+m} &= H\left(\bigvee_{i=0}^{n+m-1} T^{-i}(\mathcal{P})\right) \\
&= H\left(\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) \vee T^{-n}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})\right)\right) \\
&\leq H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) + H\left(T^{-n}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})\right)\right) \\
&= u_n + u_m.
\end{aligned}$$

Ceci implique que la suite $(u_n/n)_{n \geq 1}$ est convergente et que sa limite est comprise entre 0 et $u_1$. En effet, fixons $p \geq 1$. Pour tout $n \geq p$, notons $n = kp + r$ la division euclidienne de $n$ par $p$ et remarquons que

$$\frac{u_n}{n} \leq \frac{u_{kp} + u_r}{n} \leq \frac{u_{kp}}{kp} + \frac{u_r}{n} \leq \frac{u_p}{p} + \frac{u_r}{n}.$$

Ceci implique que

$$\limsup_{n\to+\infty} \frac{u_n}{n} \leq \frac{u_p}{p} \quad (\leq u_1),$$

et donc que

$$\limsup_{n\to+\infty} \frac{u_n}{n} \leq \inf_{p\geq 1} \frac{u_p}{p} \leq \liminf_{n\to+\infty} \frac{u_n}{n}.$$

□

L'entropie $h(T, \mathcal{P})$ peut être définie de la façon équivalente suivante :

Proposition 4.3.2 : *Si $\mathcal{P}$ est une partition mesurable de $X$, alors la suite*

$$H\left(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right)$$

*est décroissante et converge vers $h(T, \mathcal{P})$.*

*Démonstration.* La décroissance de cette suite est une conséquence de l'assertion **v)** de la proposition 4.2.1. Notons $l$ sa limite. Remarquons que

$$\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) = \mathcal{P} \vee \left(\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right),$$

ce qui implique

$$\begin{aligned}
H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) &= H\left(\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right) + H\left(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right) \\
&= H\left(\bigvee_{i=0}^{n-2} T^{-i}(\mathcal{P})\right) + H\left(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right) \\
&= H(\mathcal{P}) + \sum_{k=1}^{n-1} H\left(\mathcal{P}|\bigvee_{i=1}^{k} T^{-i}(\mathcal{P})\right).
\end{aligned}$$

Par le théorème de Césaro, on obtient

$$h(T, \mathcal{P}) = \lim_{n\to+\infty} H\left(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right).$$

□

Énonçons les propriétés principales de l'entropie relativement aux partitions.

Proposition 4.3.3 : *Soient $\mathcal{P}$ et $\mathcal{Q}$ des partitions mesurables. On a alors les propriétés suivantes :*

**i)** $h(T, \mathcal{P}) \leq h(T, \mathcal{Q}) + H(\mathcal{P}|\mathcal{Q})$ ;

**ii)** $h(T, \mathcal{Q}) \leq h(T, \mathcal{P})$ *si* $\mathcal{Q} \preceq \mathcal{P}$ ;

**iii)** $h(T, \bigvee_{i=0}^{m} T^{-i}(\mathcal{P})) = h(T, \mathcal{P})$ *pour tout* $m \geq 0$ ;

**iv)** $h(T^m, \bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})) = mh(T, \mathcal{P})$ *pour tout* $m \geq 1$ ;

**v)** $h(T, T^{-1}(\mathcal{P})) = h(T, \mathcal{P})$ ;

**v)** $h(T^{-1},\mathcal{P}) = h(T,\mathcal{P})$ *si* $T$ *est inversible.*

*Démonstration.* Pour prouver **i)**, remarquons que

$$\begin{aligned}
H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) &\leq H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) | \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right)\\
&\leq H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + \sum_{i=0}^{n-1} H\left(T^{-i}(\mathcal{P}) | \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right)\\
&\leq H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + \sum_{i=0}^{n-1} H(T^{-i}(\mathcal{P})|T^{-i}(\mathcal{Q}))\\
&= H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + nH(\mathcal{P}|\mathcal{Q}).
\end{aligned}$$

Il reste à diviser par $n$ et à faire tendre $n$ vers l'infini pour obtenir **i)**.

On en déduit **ii)** puisque $H(\mathcal{Q}|\mathcal{P}) = 0$ si $\mathcal{Q} \preceq \mathcal{P}$. On peut également déduire **ii)** des relations

$$\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q}) \preceq \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}).$$

Pour obtenir **iii)** posons $\mathcal{S} = \bigvee_{i=0}^{m} T^{-i}(\mathcal{P})$ et remarquons qu'on a

$$\bigvee_{i=0}^{m+n-1} T^{-i}(\mathcal{P}) \sim \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{S}),$$

ce qui implique

$$\begin{aligned}
h(T,\mathcal{P}) &= \lim_{n\to+\infty} \frac{1}{n+m} H(\left(\bigvee_{i=0}^{m+n-1} T^{-i}(\mathcal{P})\right)\\
&= \lim_{n\to+\infty} \frac{n}{n+m}\frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{S})\right) = h(T,\mathcal{S}).
\end{aligned}$$

Pour établir **iv)**, remarquons qu'on a

$$\bigvee_{i=0}^{n-1} T^{-im}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})\right) = \bigvee_{i=0}^{nm-1} T^{-i}(\mathcal{P}),$$

ce qui implique

$$\begin{aligned}
h(T^m, \bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})) &= \lim_{n\to+\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-im}\left(\bigvee_{i=0}^{m-1} T^{-i}(\mathcal{P})\right)\right)\\
&= \lim_{n\to+\infty} \frac{m}{nm} H\left(\bigvee_{i=0}^{nm-1} T^{-i}(\mathcal{P})\right) = mh(T,\mathcal{P}).
\end{aligned}$$

L'assertion **v)** se déduit de

$$H\left(\bigvee_{i=0}^{n-1} T^{-i}(T^{-1}(\mathcal{P}))\right) = H\left(T^{-1}\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)\right) = H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right),$$

et l'assertion **vi)** de

$$H\left(\bigvee_{i=0}^{n-1} T^i(\mathcal{P})\right) = H\left(T^n\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)\right) = H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right).$$

□

On définit alors l'entropie métrique $h(T) \in [0, +\infty]$ comme le supremum des entropies relativement à toutes les partitions :

$$h(T) = \sup\{h(T, \mathcal{P}) \ , \ \mathcal{P} \text{ est une partition mesurable de } X\}.$$

Proposition 4.3.4 : *On a les propriétés suivantes :*

**i)** $h(T^n) = nh(T)$, *pour tout* $n \geq 0$ *;*

**ii)** *si* $T$ *est inversible, alors* $h(T^k) = |k|h(T)$, *pour tout* $k \in \mathbf{Z}$.

*Démonstration.* L'assertion **i)** est évidemment vraie si $n = 0$. En effet, pour toute partition mesurable $\mathcal{P}$ et tout $n \geq 1$, on a $\bigvee_{i=0}^{n-1} \mathrm{Id}_X^{-i}(\mathcal{P}) \sim \mathcal{P}$, ce qui implique que

$$h(\mathrm{Id}_X, \mathcal{P}) = \lim_{n\to+\infty} \frac{1}{n} H(\mathcal{P}) = 0.$$

Supposons donc $n \geq 1$ et remarquons que pour toute partition mesurable $\mathcal{P}$, on a

$$h(T^n, \mathcal{P}) \leq h\left(T^n, \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) = nh(T, \mathcal{P}),$$

ce qui nous donne

$$h(T^n, \mathcal{P}) \leq nh_\mu(T), \ \ nh(T, \mathcal{P}) \leq h_\mu(T^n)$$

puis

$$h(T^n) \leq nh_\mu(T), \ \ nh(T) \leq h_\mu(T^n).$$

Pour obtenir **ii)**, il suffit de remarquer que pour toute partition mesurable $\mathcal{P}$, on a $h(T^n, \mathcal{P}) = h(T^{-n}, \mathcal{P})$.

□

Nous allons prouver maintenant que l'entropie est un invariant de conjugaison. Si l'on veut spécifier la mesure $\mu$, on peut écrire $H_\mu(\mathcal{P})$ pour l'entropie d'une partition mesurable $\mathcal{P}$ ; si $T : X \to X$ est mesurable et préserve $\mu$, on peut écrire $h_\mu(T, \mathcal{P})$ pour l'entropie relativement à $\mathcal{P}$ et $h_\mu(T)$ pour l'entropie.

Proposition 4.3.5 : *Soient* $(X, \mathcal{B}, \mu, T)$ *et* $(Y, \mathcal{C}, \nu, S)$ *deux systèmes dynamiques mesurés, tels que* $\mu(X) = \nu(Y) = 1$. *Supposons qu'il existe une application mesurable* $R\ : X \to Y$ *telle que* $R \circ T = S \circ R$ *et* $R_*(\mu) = \nu$. *Alors* $h_\nu(S) \leq h_\mu(T)$. *De plus il y a égalité si* $R$ *est inversible et* $R^{-1}$ *est mesurable.*

*Démonstration.* Pour toute partition mesurable $\mathcal{P}$ de $Y$, et tout entier $n \geq 0$, on a

$$\bigvee_{i=0}^{n-1} T^{-i}(R^{-1}(\mathcal{P})) = R^{-1}\left(\bigvee_{i=0}^{n-1} S^{-i}(\mathcal{P})\right),$$

ce qui implique

$$H_\mu\left(\bigvee_{i=0}^{n-1} T^{-i}(R^{-1}(\mathcal{P}))\right) = H_\nu\left(\bigvee_{i=0}^{n-1} S^{-i}(\mathcal{P})\right).$$

On en déduit

$$h_\nu(S,\mathcal{P}) = h_\mu(T, R^{-1}(\mathcal{P})) \leq h_\mu(T),$$

et donc $h_\nu(S) \leq h_\mu(T)$. On a bien sûr égalité si $R$ est inversible et $R^{-1}$ est mesurable. □

L'inégalité peut bien sûr être stricte. L'exemple le plus simple est le cas où $\nu$ est la mesurable de Dirac en un point fixe $y_0 \in Y$ de $S$ et $R$ l'application constante égale à $y_0$. Dans ce cas on a $h_\nu(S) = 0$. En fait, $h_\nu(S)$ est le supremum des entropies (pour $T$) parmi les partitions mesurables qui sont images d'une partition de $Y$. Dans l'exemple précédent il n'y a qu'une seule partition de ce type, la partition triviale $\{X\}$. Notons également que dans la seconde assertion, l'inversibilité n'est nécessaire que presque partout : il existe deux ensembles mesurables $X' \subset X$ et $Y' \subset Y$ tels que $\mu(X') = \nu(Y') = 1$ et tels que $R|_{X'}$ est une bijection de $X'$ sur $Y'$ dont l'inverse est mesurable.

## 4.4 Partitions génératrices

Nous verrons dans cette courte section comment, grâce à Kolmogorov et Sinai, il suffit parfois d'une partition pour définir l'entropie.

Proposition 4.4.1 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. Si $(\mathcal{P}_m)_{m\geq 0}$ est une suite croissante génératrice de partitions mesurables, alors*

$$h(T) = \lim_{m\to+\infty} h(T, \mathcal{P}_m).$$

*Démonstration.* La suite $(h(T,\mathcal{P}_m))_{m\geq 0}$ est croissante et converge donc dans $[0,+\infty]$. Pour toute partition mesurable $\mathcal{P}$ et tout $\varepsilon > 0$, nous avons vu qu'il était possible de trouver un entier $m_0$ et une partition $\mathcal{Q}$ vérifiant $\mathcal{Q} \preceq \mathcal{P}_{m_0}$ et $H(\mathcal{P}|\mathcal{Q}) \leq \varepsilon$. On en déduit

$$h(T,\mathcal{P}) \leq h(T,\mathcal{Q}) + H(\mathcal{P}|\mathcal{Q}) \leq h(T,\mathcal{P}_m) + \varepsilon \leq \lim_{m\to+\infty} h(T,\mathcal{P}_m) + \varepsilon,$$

puis

$$h(T) \leq \lim_{m\to+\infty} h(T,\mathcal{P}_m).$$

□

**Définition.** Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$. On dira qu'une partition $\mathcal{P}$ est *fortement génératrice* si la suite $\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right)_{n\geq 1}$ est génératrice. Dans le cas où $T$ est inversible, on dira que $\mathcal{P}$ est*génératrice* si la suite $\left(\bigvee_{i=-n+1}^{n-1} T^{-i}(\mathcal{P})\right)_{n\geq 1}$ est génératrice.

Théorème 4.4.2 : *Soit $(X, \mathcal{B}, \mu, T)$ un système dynamique mesuré tel que $\mu(X) = 1$ et $\mathcal{P}$ une partition mesurable. Si $\mathcal{P}$ est fortement génératrice, alors $h_\mu(T) = h(T,\mathcal{P})$. On a la même égalité si $T$ est inversible et si $\mathcal{P}$ est génératrice.*

*Démonstration.* Nous allons utiliser la proposition 4.4.1 ainsi que l'assertion **iii)** de la proposition 4.3.3. Écrivons

$$h(T) = \lim_{n\to+\infty} h\left(T, \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) = \lim_{n\to+\infty} h(T,\mathcal{P}) = h(T,\mathcal{P}).$$

Dans le cas où $T$ est inversible et $\mathcal{P}$ est un générateur, écrivons

$$h(T) = \lim_{n\to+\infty} h\left(T, \bigvee_{i=-n+1}^{n-1} T^{-i}(\mathcal{P})\right) = \lim_{n\to+\infty} h\left(T, \bigvee_{i=0}^{2n-2} T^{-i}(\mathcal{P})\right) = h(T,\mathcal{P}).$$

□

## 4.5 Exemples

Nous allons calculer des entropies à l'aide notamment du théorème 4.4.2.

**Le décalage de Bernouilli**

Considérons le décalage $\sigma$ sur $\{1,\dots,r\}^{\mathbf{N}}$. Remarquons que la (vraie) partition $\mathcal{P} = (P_i)_{1\leq i\leq r}$ où $P_i = \{x = (x_n)_{n\geq 0} \,|\, x_0 = i\}$ est fortement génératrice car $\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{P})$ est la partition en les $p^n$ cylindres de la forme $C^0_{(w_0,\dots,w_{n-1})}$. Pour toute mesure invariante $\mu$, on a donc l'égalité

$$h_\mu(\sigma) = h_\mu(\sigma,\mathcal{P}).$$

Si $\mu$ est une mesure produit définie par une famille $(p_i)_{1\leq i\leq r}$ de réels positifs telle que $\sum_{i=1}^r p_i = 1$, alors, pour tout $n\geq 0$ on a

$$\begin{aligned} H(\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{P})) &= \sum_{i_0,\dots,i_{n-1}} -p_{i_0}\dots p_{i_{n-1}} \ln(p_{i_0}\dots p_{i_{n-1}}) \\ &= n\sum_{i=1}^r p_i \ln p_i. \end{aligned}$$

On en déduit que

$$h_\mu(\sigma) = \sum_{i=1}^r -p_i \ln p_i = H(\mathcal{P}).$$

**Remarques**

1. On a $h_\mu(\sigma) \leq \ln r$, avec égalité uniquement dans le cas de la mesure équidistribuée (voir également exercice **)

2. Si $r\neq s$, les décalages sur $\{1,\dots,r\}^N$ et $\{1,\dots,s\}^N$, ne sont pas conjugués, considérés comme systèmes dynamiques mesurés avec la mesure équidistribuée, car ils n'ont pas la même entropie.

Supposons maintenant que $\mu$ est une mesure de Markov définie par une matrice stochastique $M = (M_{i,j})_{i,j}$ et par un vecteur fixe $v = (v_1,\dots,v_r)$ à coefficients positifs tels que $\sum_{i=1}^r v_i = 1$. Pour tout $w = (w_0,\dots,w_{m-1}) \in \{1,\dots,r\}^m$ et tout $n_0\in\mathbf{N}$, on a

$$\mu(C_w^{n_0}) = v_{w_0}\left(\prod_{k=0}^{m-2} M_{w_k,w_{k+1}}\right),$$

On sait que

$$h_\mu(\sigma) = h_\mu(\sigma, \mathcal{P}) = \lim_{n\to+\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{P})).$$

Le calcul donne alors

$$\begin{aligned}
H(\bigvee_{i=0}^{n-1} \sigma^{-i}(\mathcal{P})) &= -\sum_{i_0,i_1,\ldots,i_{n-1}} \mu(C^0_{i_0,i_1,\ldots,i_{n-1}}) \ln\Big(\mu(C^0_{i_0,i_1,\ldots,i_{n-1}})\Big) \\
&= -\sum_{i_0,i_1,\ldots,i_{n-1}} v_{i_0} M_{i_0,i_1} \ldots M_{i_{n-2},i_{n-1}} \ln(v_{i_0} M_{i_0,i_1} \ldots M_{i_{n-2},i_{n-1}}) \\
&= -\sum_{i_0,i_1,\ldots,i_{n-1}} v_{i_0} M_{i_0,i_1} \ldots M_{i_{n-2},i_{n-1}} (\ln v_{i_0} + \ln M_{i_0,i_1} + \ldots + \ln M_{i_{n-2},i_{n-1}}) \\
&= -\sum_{i_0} v_{i_0} \ln v_{i_0} - (n-1) \sum_{i,j} v_i M_{i,j} \ln M_{i,j}
\end{aligned},$$

et donc

$$h_\mu(\sigma) = -\sum_{i,j} v_i M_{i,j} \ln M_{i,j}.$$

**Remarques**

1. Ce qui précède peut être fait sur le décalage bilatéral. Écrivons $P_i = \{x = (x_k)_{k\in\mathbf{Z}} \mid x_0 = i\}$. La partition $\mathcal{P} = (P_i)_{1\le i\le r}$ étant génératrice (mais pas fortement génératrice) on a, pour toute mesure invariante $\mu$, l'égalité

$$h_\mu(\sigma) = h_\mu(\sigma, \mathcal{P}).$$

Pour les mesures produits, on trouvera les mêmes formules que dans le cas unilatéral

2. Un difficile résultat d'Ornstein exprime que deux décalages bilatéraux, considérés comme systèmes dynamiques mesurés avec chacun une mesure de type produit, sont conjugués si et seulement s'ils ont même entropie (le résultat est faux pour les décalages unilatéraux).

**Endomorphismes linéaires de $\mathbf{T}^1$.**

L'application $F\ :\ \widehat{x} \to p\widehat{x}$ sur $\mathbf{T}^1$, où $|p| \geq 2$, préserve la mesure de Haar. Nous allons calculer l'entropie du système dynamique mesuré obtenu. On peut se limiter au cas où $p \geq 2$, puisque $h(F^2) = 2h(F)$ et que $F^2(\widehat{x}) = p^2\widehat{x}$. Or nous avons vu que le système est alors conjugué au décalage de Bernouilli sur $\{1, \ldots, p\}$ avec la mesure équidistribuée. On a donc généralement l'égalité $h(F) = \ln|p|$. On aurait également pu utiliser directement la partition fortement génératrice $\mathcal{P} = (P_i)_{0\le i\le p-1}$, où $P_i$ est la projection dans $\mathbf{T}^1$ de $[i/p, (i+1)/p[$.

**Rotations du cercle.**

La rotation $T_{\widehat{a}}\ :\ \widehat{x} \to \widehat{x} + \widehat{a}$ préserve la mesure de Haar sur $\mathbf{T}^1$. Si $\widehat{a} \in \mathbf{Q}/\mathbf{Z}$, il existe $q \geq 1$ tel que $T^q_{\widehat{a}} = \mathrm{Id}_{\mathbf{T}^1}$ et donc on a $h(T_{\widehat{a}}) = 0$. Si $\alpha \notin \mathbf{Q}/\mathbf{Z}$, on peut vérifier que la partition $\mathcal{P} = (P_0, P_1)$ est fortement génératrice, où $P_i$ est la projection dans $\mathbf{T}^1$ de $[i/2, (i+1)/2[$. Ceci implique que $h_\mu(T_{\widehat{a}}) = 0$. En effet $T^{-1}(\mathcal{P})$ est également fortement génératrice, et donc :

$$h(T) = h(T, \mathcal{P}) = \lim_{n\to+\infty} H(\mathcal{P} | \bigvee_{i=1}^{n} T^{-i}(\mathcal{P})) = 0.$$

Remarquons que nous avons le résultat plus général suivant :

Proposition 4.5.1 : *Soit* $(X, \mathcal{B}, \mu, T)$ *un système dynamique mesuré inversible tel que* $\mu(X) = 1$*. S'il existe une partition fortement génératrice, alors* $h(T) = 0$.

## 5.1 Le tore $\mathbf{T}^r$.

On notera $\mathbf{T}^r$ le groupe quotient $\mathbf{T}^r = \mathbf{R}^r/\mathbf{Z}^r$ et

$$\pi : \mathbf{R}^r \to \mathbf{T}^r, \quad x \mapsto x + \mathbf{Z}^r$$

la projection. On munit $\mathbf{T}^r$ de la topologie quotient, formée des parties $U$ dont la préimage $\pi^{-1}(U)$ est ouverte dans $\mathbf{R}^r$. Cette topologie est caractérisée par la propriété suivante :

PROPOSITION 5.1.1 : *Soit $X$ un espace topologique. Une fonction $F\ : \mathbf{T}^r \to X$ est continue si et seulement si $F \circ \pi\ : \mathbf{R}^r \to X$ est continue.*

Cela implique :

COROLLAIRE 5.1.2 : *Soit $X$ un espace topologique. Si une fonction continue $f\ : \mathbf{R}^r \to X$ est invariante par les translations entières, alors il existe une application continue $F\ : \mathbf{T}^r \to X$ telle que $f = F \circ \pi$.*

PROPOSITION 5.1.3 : *Le tore $\mathbf{T}^r$ est un espace connexe et compact homéomorphe au produit de $r$ cercles euclidiens $S^1$, où $S^1 = \{(a,b) \in \mathbf{R}^2 \,|\, a^2 + b^2 = 1\}$.*

*Preuve.* Remarquons d'abord que $\mathbf{T}^r$, muni de la topologie quotient est séparé. Remarquons ensuite que $\mathbf{T}^r = \pi([0,1]^r)$. Puisque $\pi$ est continue, on en déduit que $\mathbf{T}^r$ est compact et connexe. Définissons

$$h : (x_1, \ldots, x_r) \mapsto (\cos(2\pi x_1), \sin(2\pi x_1), \ldots, \cos(2\pi x_r), \sin(2\pi x_r)).$$

C'est une application continue de $\mathbf{R}^r$ dans $(S^1)^r$. Elle est invariante par les translation entières et définit une application continue $H\ : \mathbf{T}^r \to (S^1)^r$ qui est en fait bijective. Puisque $\mathbf{T}^n$ est compact, on en déduit que $F$ est un homéomorphisme. $\square$

On peut définir la topologie quotient par des distances naturelles. Si $\|\ \|$ est une norme sur $\mathbf{R}^r$, on peut définir pour tous $\widehat{x}$, $\widehat{y}$ dans $\mathbf{T}^r$ :

$$d(\widehat{x}, \widehat{y}) = \inf_{\pi(x)=\widehat{x},\, \pi(y)=\widehat{y}} \|x - y\|.$$

Fixons $x_0$ et $y_0$ tels que $\pi(x_0) = \widehat{x}$ et $\pi(y_0) = \widehat{y}$. Remarquons que

$$\begin{aligned} d(\widehat{x}, \widehat{y}) = \inf_{\pi(y)=\widehat{y}} \|x_0 - y\| &= \inf_{k \in \mathbf{Z}^n} \|x_0 - y_0 - k\| \\ &= \min_{k \in \mathbf{Z}^n} \|x_0 - y_0 - k\| = \min_{\pi(x)=\widehat{x},\, \pi(y)=\widehat{y}} \|x - y\| \end{aligned}$$

car $\lim_{\|k\| \to +\infty} \|x_0 - y_0 - k\| = +\infty$.

PROPOSITION 5.1.4 : *On a les propriétés suivantes :*

**i)** *l'application $d$ est une distance sur $\mathbf{T}^r$ qui induit la topologie quotient ;*

**ii)** *si $d'$ est définie similairement à partir d'une norme $\| \ \|'$ sur $\mathbf{R}^r$, alors $d$ et $d'$ sont équivalentes ;*

**iii)** *il existe $\varepsilon > 0$, tel que pour tout $\widehat{x} \in \mathbf{T}^r$, on a*

$$\pi^{-1}(B(\widehat{x}, \varepsilon)) = \bigsqcup_{\pi(x)=\widehat{x}} B(x, \varepsilon),$$

*et chaque application $\pi|_{B(x,\varepsilon)}$ est une isométrie de $B(x, \varepsilon)$ sur $B(\widehat{x}, \varepsilon)$.*

*Preuve.* Prouvons que $d$ est une distance. L'égalité $d(\widehat{x}, \widehat{y}) = 0$ signifie qu'il existe $x \in \pi^{-1}(\{\widehat{x}\})$ et $y \in \pi^{-1}(\{\widehat{y}\})$ tels que $\|x - y\| = 0$. Cela implique que $x = y$ et donc que $\widehat{x} = \widehat{y}$.

Le caractère symétrique de $d$ étant évident, il reste à prouver

$$d(\widehat{x}, \widehat{z}) \leq d(\widehat{x}, \widehat{y}) + d(\widehat{y}, \widehat{z}).$$

Fixons $y \in \pi^{-1}(\{\widehat{y}\})$. Nous savons qu'il existe $x \in \pi^{-1}(\{\widehat{x}\})$ et $z \in \pi^{-1}(\{\widehat{z}\})$ tels que $d(\widehat{x}, \widehat{y}) = \|x - y\|$ et $d(\widehat{y}, \widehat{z}) = \|y - z\|$. On en déduit

$$d(\widehat{x}, \widehat{z}) \leq \|x - z\| \leq \|x - y\| + \|y - z\| = d(\widehat{x}, \widehat{y}) + \widehat{d}(\widehat{y}, \widehat{z}).$$

Prouvons l'assertion **ii)**. Les deux normes $\| \ \|$ et $\| \ \|'$ étant équivalentes, il existe $C_1 > 0$ et $C_2 > 0$ tels que pour tout $x \in \mathbf{R}^r$, on a

$$C_1\|x\| \leq \|x\|' \leq C_2\|x\|.$$

Pour tous $\widehat{x}$ et $\widehat{y}$ dans $\mathbf{T}^r$, il existe $x \in \pi^{-1}(\{\widehat{x}\})$ et $y \in \pi^{-1}(\{\widehat{y}\})$ tels que $d(\widehat{x}, \widehat{y}) = \|x - y\|$. On en déduit

$$d'(\widehat{x}, \widehat{y}) \leq \|x - y\|' \leq C_2\|x - y\| = C_2 d(\widehat{x}, \widehat{y}).$$

On établit de même

$$C_1 d(\widehat{x}, \widehat{y}) \leq d'(\widehat{x}, \widehat{y}).$$

Les distances $d$ et $d'$ sont donc équivalentes.

Démontrons maintenant l'assertion **iii)**. Définissons $\varepsilon$ en posant

$$\inf_{k \in \mathbf{Z}^r \setminus \{0\}} \|k\| = \min_{k \in \mathbf{Z}^r \setminus \{0\}} \|k\| = 4\varepsilon > 0.$$

Remarquons que pour tout $x \in \mathbf{R}^r$, les boules ouvertes $B(x + k, \varepsilon)$, $k \in \mathbf{Z}^r$, sont disjointes deux-à-deux. La réunion de ces boules n'est rien d'autre que $\pi^{-1}(B(\pi(x), \varepsilon))$. On doit montrer que $\pi|_{B(x,\varepsilon)}$ est une isométrie de $B(x, \varepsilon)$ sur $B(\pi(x), \varepsilon)$. Le fait que les boules $B(x+k, \varepsilon)$, $k \in \mathbf{Z}^r$, sont disjointes deux-à-deux implique que tout point $\widehat{y} \in B(\pi(x), \varepsilon)$ a un antécédent par $\pi$ dans $B(x, \varepsilon)$ et que cet antécédent $y$ est unique. De plus, puisque $\|x - y\| < \|x - y - k\|$ pour tout $k \in \mathbf{Z}^r \setminus \{0\}$, on sait que $\|x - y\| = d(\pi(x), \widehat{y})$.

Il reste à prouver que $d$ induit la topologie quotient sur $\mathbf{T}^r$. L'application $\pi$ étant 1-lipschitzienne, elle est continue, ce qui prouve que la topologie induite par $d$ est moins fine que la topologie quotient. Puisque $\mathbf{T}^r$, muni de la topologie induite par $d$ est séparé, et muni des la topologie quotient est compact, on en déduit que les deux topologies coïncident. Pour montrer qu'elle est plus fine, il reste à prouver que si $X$ est un espace topologique, alors toute application

$F : \mathbf{T}^r \to X$ telle que $F \circ \pi$ est continue, doit être continue pour la topologie induite par $d$. Pour tout $\widehat{x} \in \mathbf{T}^r$, on peut choisir $x \in \pi^{-1}(\{\widehat{x}\})$ et écrire

$$F|_{B(\widehat{x},\varepsilon)} = (F \circ \pi) \circ (\pi|_{B(x,\varepsilon)})^{-1}$$

comme composition d'applications continues. Ceci implique bien sûr que $F$ est continue en $\widehat{x}$.□

Si $f : \mathbf{R}^r \to \mathbf{R}^r$ est une application continue telle que

$$x - y \in \mathbf{Z}^r \Rightarrow f(x) - f(y) \in \mathbf{Z}^r,$$

alors $\pi \circ f : \mathbf{R}^r \to \mathbf{T}^r$ est continue et invariante par les translations entières. On en déduit qu'il existe une application continue $F : \mathbf{T}^r \to \mathbf{T}^r$ telle que $\pi \circ f = F \circ \pi$. On dit que $f$ *relève* $F$.

Remarquons que si $f$ relève $F$ et $g$ relève $G$, alors $f \circ g$ relève $F \circ G$, puisque

$$\pi \circ f \circ g = F \circ \pi \circ g = F \circ G \circ \pi.$$

Cela implique que $f^n$ relève $F^n$, pour tout $n \geq 0$. Notons aussi que $f + g$ relève $F + G$ puisque

$$\pi \circ (f + g) = \pi \circ f + \pi \circ g = F \circ \pi + G \circ \pi = (F + G) \circ \pi.$$

PROPOSITION 5.1.5 : *Si $f$ relève $F$, les autres relèvements de $F$ sont les applications $T_k \circ f$, $k \in \mathbf{Z}^r$, où $T_k : x \mapsto x + k$.*

*Preuve.* Si $g(x) = f(x) + k$, $k \in \mathbf{Z}$, on a $\pi \circ g(x) = \pi \circ f(x) = F \circ \pi(x)$, ainsi $g$ relève $F$.

Supposons maintenant que $g$ relève $F$. Puisque $g - f$ relève $F - F = 0$, on sait que $g(x) - f(x) \in \mathbf{Z}^r$ pour tout $x \in \mathbf{R}^r$. Puisque $g - f$ est continue, c'est une application constante : il existe $k \in \mathbf{Z}^r$ tel que $g(x) - f(x) = k$. □

Rappelons le résultat fondamental suivant, conséquence du fait que $\mathbf{R}^r$ est le revêtement universel de $\mathbf{T}^r$ et $\pi$ la projection de revêtement.

THÉORÈME 5.1.6 : *Toute application continue $F : \mathbf{T}^r \to \mathbf{T}^r$ admet un relèvement.*

On en déduit :

COROLLAIRE 5.1.7 : *Si $F$ est un homéomorphisme de $\mathbf{T}^r$, alors tout relèvement $f$ de $F$ est un homéomorphisme de $\mathbf{R}^r$ et $f^{-1}$ relève $F^{-1}$.*

*Preuve.* Soit $g$ un relèvement de $F^{-1}$. Alors $f \circ g$ et $g \circ f$ relèvent $\mathrm{Id}_{\mathbf{T}^r}$. Ainsi, il existe $k \in \mathbf{Z}^r$ et $k' \in \mathbf{Z}^r$ tels que $f \circ g = T_k$ et $g \circ f = T_{k'}$. La première égalité implique que $f$ est surjective et la seconde que $f$ est injective. Puisque $f$ est bijective, on peut écrire $f^{-1} = T_{k'}^{-1} \circ g$, ce qui implique que $f$ est un homéomorphisme et que $f^{-1}$ relève $F^{-1}$. □

**Remarques.**

1. Une application non injective $F : \mathbf{T}^r \to \mathbf{T}^r$ peut être relevée par un homéomorphisme $f$. C'est le cas par exemple de $f : \mathbf{R} \to \mathbf{R}, x \mapsto 2x$ qui relève $F : \mathbf{T}^1 \to \mathbf{T}^1, \widehat{x} \to 2\widehat{x}$, et on a $F(\pi(1/2)) = F(\pi(0)) = \pi(0)$. On prouve facilement qu'un homéomorphisme $f : \mathbf{R}^r \to \mathbf{R}^r$ relève un homéomorphisme $\mathbf{T}^r$ si et seulement si:

- $f(x + k) - f(x) \in \mathbf{Z}^r$, pour tout $x \in \mathbf{R}^r$ et tout $k \in \mathbf{Z}^r$;

- $f^{-1}(x+k) - f^{-1}(x) \in \mathbf{Z}^n$, pour tout $x \in \mathbf{R}^r$ et tout $k \in \mathbf{Z}^r$.

2. Si une application continue $F : \mathbf{T}^r \to \mathbf{T}^r$ est relevée par une application différentiable $f$, alors tous les relèvements ont la même propriété et on peut dire que $F$ est *différentiable.* De façon analogue, si $F$ est un homéomorphisme et $f$ un difféomorphisme, on dira que $F$ est un *difféomorphisme.* On peut définir de même des applications *de classe $C^p$* ou*réelles analytiques* sur $\mathbf{T}^r$. Par exemple

$$f : (x_1, x_2) \mapsto (x_1 + x_2,\, x_2 + \sin(2\pi(x_1 + x_2)))$$

relève un difféomorphisme réel analytique $F$.

## 5.2 Rotations sur $\mathbf{T}^r$, mesure de Haar.

Rappelons dans cette section un résultat classique, en notant $T_{\widehat{a}} : \widehat{x} \mapsto \widehat{x} + \widehat{a}$ la rotation d'angle $\widehat{a} \in \mathbf{T}^r$, qui est relevée par les translations $T_a : x \mapsto x + a$, $a \in \pi^{-1}(\{\widehat{a}\})$.

Théorème 5.2.1 : *Il existe une unique mesure borélienne de probabilité $\mu$, invariante par toute rotation $T_{\widehat{a}}$. C'est la* mesure de Haar *du groupe topologique $\mathbf{T}^r$. Pour toute application continue $\varphi : \mathbf{T}^r \to \mathbf{C}$ on a*

$$\int \varphi \, d\mu = \int_{[0,1]^r} \varphi \circ \pi \, d\lambda,$$

*où $\lambda$ est la mesure de Lebesgue sur $\mathbf{R}^r$.*

*Preuve.* Notons $C(\mathbf{T}^r, \mathbf{C})$ l'espace vectoriel des applications continues $\varphi : \mathbf{T}^r \to \mathbf{C}$, muni de la norme

$$\|\varphi\| = \sup_{\widehat{x} \in \mathbf{T}^r} |\varphi(\widehat{x})|.$$

L'application

$$L : \varphi \mapsto \int_{[0,1]^n} \varphi \circ \pi \, d\lambda$$

est une forme linéaire continue sur $C(\mathbf{T}^r, \mathbf{C})$ qui est positive ($\varphi \geq 0 \Rightarrow L(\varphi) \geq 0$) et qui envoie la fonction constante égale à 1 sur 1. Par le théorème de représentation de Riesz, on sait qu'il existe une mesure borélienne de probabilité $\mu$ sur $\mathbf{T}^r$ telle que pour toute fonction continue $\varphi : \mathbf{T}^r \to \mathbf{C}$, on a $L(\varphi) = \int_{\mathbf{T}^r} \varphi \, d\mu$.

Montrons que $\mu$ est invariante par toute rotation $T_{\widehat{a}}$. Fixons $a \in \widehat{a}$ et écrivons :

$$\begin{aligned}\int_{\mathbf{T}^r} \varphi \circ T_{\widehat{a}} \, d\mu = \int_{[0,1]^n} \varphi \circ T_{\widehat{a}} \circ \pi \, d\lambda &= \int_{[0,1]^r} \varphi \circ \pi \circ T_a \, d\lambda \\ &= \int_{[0,1]^r + a} \varphi \circ \pi \, d\lambda = \int_{[0,1]^r} \varphi \circ \pi \, d\lambda = \int_{\mathbf{T}^r} \varphi \, d\mu,\end{aligned}$$

puisque la fonction $\varphi \circ \pi$ est invariante par toute translation entière et la mesure $\lambda$ est invariante par toute translation.

Il reste à prouver qu'elle est unique. Si $\mu'$ est également invariante par toute rotation, il faut montrer que $\int \varphi \, d\mu = \int \varphi \, d\mu'$ pour toute application $\varphi \in C(\mathbf{T}^r, \mathbf{C})$. Fixons un entier $m \geq 0$, des entiers $k_1, \ldots, k_r$ in $\{0, \ldots, 2^m - 1\}$, et posons

$$C^m_{k_1,\ldots,k_r} = \pi\left(\left[\frac{k_1}{2^m}, \frac{k_1+1}{2^m}\right[ \times \ldots \times \left[\frac{k_r}{2^m}, \frac{k_r+1}{2^m}\right[\right).$$

Écrivons $\chi_{C^m_{k_1,\dots,k_r}}$ pour la fonction caractéristique de $C^m_{k_1,\dots,k_r}$.

On a une partition $\mathbf{T}^r$ en $2^{rm}$ ensembles boréliens mutuellement translatés, on en déduit que

$$\mu\left(C^m_{k_1,\dots,k_r}\right) = \mu'\left(C^m_{k_1,\dots,k_r}\right) = \frac{1}{2^{rm}}.$$

Toute fonction $\varphi \in C(\mathbf{T}^r, \mathbf{C})$, étant uniformément continue, est uniformément approchée par une suite $(\varphi_m)_{m\geq 0}$, où $\varphi_m = \sum_{k_1,\dots,k_r} \lambda^m_{k_1,\dots,k_r} \chi_{C^m_{k_1,\dots,k_r}}$. Remarquons maintenant que

$$\int \varphi_m \, d\mu = \int \varphi_m \, d\mu' = \frac{1}{2^{rm}} \sum_{k_1,\dots,k_r} \lambda^m_{k_1,\dots,k_r}$$

pour tout $m \geq 0$. On déduit que

$$\int \varphi \, d\mu = \lim_{m\to+\infty} \int \varphi_m \, d\mu = \lim_{m\to+\infty} \int \varphi_m \, d\mu' = \int \varphi \, d\mu'.$$

$\square$

### 5.3 Endomorphismes linéaires de $\mathbf{T}^r$

Soit $A : \mathbf{R}^r \to \mathbf{R}^r$ une application linéaire dont la matrice dans la base canonique de $\mathbf{R}^r$ est à coefficients entiers. Puisque $A(\mathbf{Z}^r) \subset \mathbf{Z}^r$, on sait que $A(x) - A(y) \in \mathbf{Z}^r$ si $x - y \in \mathbf{Z}^r$. On en déduit que $A$ relève une application continue $\widehat{A} : \mathbf{T}^r \to \mathbf{T}^r$. Il s'agit d'un *endomorphisme linéaire*, c'est-à-dire d'un endomorphisme continu de $\mathbf{T}^r$. Il en est ainsi par exemple de

$$\begin{aligned} \widehat{A}_1 \ : \mathbf{T}^1 &\to \mathbf{T}^1 \\ \widehat{x} &\mapsto 2\widehat{x} \end{aligned}$$

et de

$$\begin{aligned} \widehat{A}_2 \ : \mathbf{T}^2 &\to \mathbf{T}^2 \\ (\widehat{x}_1, \widehat{x}_2) &\mapsto (2\widehat{x}_1 + \widehat{x}_2, \widehat{x}_1 + \widehat{x}_2) \end{aligned}.$$

PROPOSITION 5.3.1 : *L'endomorphisme $\widehat{A}$ est surjectif si et seulement si* $\det(A) \neq 0$. *Dans ce cas, il préserve la mesure de Haar. De plus, tout point a exactement* $|\det(A)|$ *antécedents. L'endomorphisme $\widehat{A}$ est bijectif si et seulement si* $\det(A) = \pm 1$. *Dans ce cas $\widehat{A}^{-1}$ est un endomorphisme linéaire.*

*Preuve.* Dire que l'application $\widehat{A}$ est surjective signifie que pour tout $y \in \mathbf{R}^r$, il existe $x \in \mathbf{R}^r$ et $k \in \mathbf{Z}^r$ tels que $y = A(x) + k$. En d'autres termes, $\widehat{A}$ est surjective si et seulement si $\mathbf{R}^r = \bigcup_{k\in\mathbf{Z}^n} T_k(\mathrm{Im}(A))$.

Si $A$ est surjective, c'est-à-dire si $\det(A) \neq 0$, alors $\widehat{A}$ l'est aussi. Si $A$ n'est pas surjective, alors l'espace quotient $\mathbf{R}^r/\mathrm{Im}(A)$ est de dimension au moins 1 et n'est donc pas dénombrable, ce qui exclut que l'on puisse écrire $\mathbf{R}^r/\mathrm{Im}(A) = \bigcup_{k\in\mathbf{Z}^r} \{k + \mathrm{Im}(A)\}$.

Pour montrer que $\widehat{A}$ préserve la mesure de Haar $\mu$ dans le cas surjectif, il suffit de prouver que la mesure $\widehat{A}_*(\mu)$ est invariante par toute rotation. On doit donc prouver que pour tout point $\widehat{b} \in \mathbf{T}^r$ et tout ensemble borélien $\widehat{X}$, on a $\mu(\widehat{A}^{-1}(T^{-1}_{\widehat{b}}(\widehat{X}))) = \mu(\widehat{A}^{-1}(\widehat{X}))$. Puisque $\widehat{A}$

est un endomorphisme surjectif, on peut trouver $\widehat{a}$ tel que $\widehat{A}(\widehat{a}) = \widehat{b}$ et on a $\widehat{A}^{-1}(T_{\widehat{b}}^{-1}(\widehat{X})) = \widehat{T}_{\widehat{a}}^{-1}(A^{-1}(\widehat{X}))$, ce qui implique que $\mu(\widehat{A}^{-1}(T_{\widehat{b}}^{-1}(\widehat{X}))) = \mu(\widehat{A}^{-1}(\widehat{X}))$ puisque $\mu$ est invariante par $T_{\widehat{a}}$.

Observons d'abord que tous les points ont le même nombre d'antécédents puisque $\widehat{A}$ est un endomorphisme. Notons $p$ ce nombre et considérons le cube $Y = [-\delta, \delta]^r$ ainsi que le polyèdre $X = A^{-1}(Y)$. Si $\delta > 0$ est petit, $Y$ se projette injectivement sur un ensemble borélien $\widehat{Y}$ de mesure $(2\delta)^r$ et $X$ sur un ensemble borélien $\widehat{X}$ de mesure $|\det(A)|^{-1}(2\delta)^r$. L'ensemble $\widehat{A}^{-1}(\widehat{Y})$ est la réunion disjointe de $p$ translatés de $\widehat{X}$. Puisque $\widehat{A}$ préserve la mesure de Lebesgue, on en déduit que $p = |\det(A)|$.

On en déduit bien sûr que $\det(A) = \pm 1$ si $\widehat{A}$ est bijectif. Dans ce cas, la matrice de $A^{-1}$ dans la base canonique est à coefficients entiers et relève un endomorphisme linéaire de $\mathbf{T}^r$ qui n'est rien d'autre que $\widehat{A}^{-1}$. $\square$

## 5.4 Classes d'homotopie des transformations continues de $\mathbf{T}^r$ dans $\mathbf{T}^r$

Intéressons-nous maintenant aux propriétés des transformations continues $F : \mathbf{T}^r \to \mathbf{T}^r$.

Proposition 5.4.1 : *Si $F : \mathbf{T}^r \to \mathbf{T}^r$ est une transformation continue, il existe une unique application linéaire $F_* : \mathbf{R}^r \to \mathbf{R}^r$ telle que pour tout relèvement $f$ de $F$, l'application $f - F_*$ est invariante par les translations entières. Sa matrice dans la base canonique est à coefficients entiers et on a $f(x+k) = f(x) + F_*(k)$, pour tout $x \in \mathbf{R}^r$ et pour tout $k \in \mathbf{Z}^r$.*

*Preuve.* Fixons un relèvement $f$ de $F$ et commençons par prouver qu'il existe au plus une application linéaire $A : \mathbf{R}^r \to \mathbf{R}^r$ telle que $f - A$ est invariante par les translations entières. Si $A_1$ et $A_2$ , vérifient cette propriété, alors $A_1 - A_2$ est invariante par les translations entières et s'annule donc sur $\mathbf{Z}^r$. On en déduit que $A_1 = A_2$.

Pour prouver l'existence, fixons $k \in \mathbf{Z}^r$ et considérons l'application $x \mapsto f(x+k) - f(x)$. Elle est continue et prend ses valeurs dans $\mathbf{Z}^r$. Il existe donc $A(k) \in \mathbf{Z}^r$ tel que $f(x+k) = f(x) + A(k)$ pour tout $x \in \mathbf{R}^r$. Remarquons que

$$\begin{aligned} A(k+k') &= f(k+k') - f(0) \\ &= f(k+k') - f(k') + f(k') - f(0) \\ &= A(k) + A(k'). \end{aligned}$$

L'homomorphisme $A : k \mapsto A(k)$ s'étend en une application linéaire $A : x \mapsto A(x)$ dont la matrice dans la base canonique est à coefficients entiers. Remarquons que $A$ ne dépend pas du choix de $f$. On peut donc écrire $A = F_*$. $\square$

Proposition 5.4.2 : *On a les propriétés suivantes :*

**i)** $(\mathrm{Id}_{\mathbf{T}^r})_* = \mathrm{Id}_{\mathbf{R}^r}$ *;*

**ii)** *$(F \circ G)_* = F_* \circ G_*$ pour toutes applications continues $F$ et $G$ ;*

**iii)** *$(F + G)_* = F_* + G_*$ pour toutes applications continues $F$ et $G$ ;*

**iv)** *si $F$ est un homéomorphisme, alors $|\det(F_*)| = 1$ et $(F^{-1})_* = (F_*)^{-1}$.*

*Preuve.* Pour prouver **i)**, il suffit de prendre $f = \mathrm{Id}_{\mathbf{R}^r}$. Pour prouver **ii)** et **iii)**, on fixe un relèvement $f$ de $F$ et un relèvement $g$ de $G$. Pour tout $x \in \mathbf{R}^r$ et tout $k \in \mathbf{Z}^r$, on a :

$$\begin{aligned} f\circ g(x+k) - f\circ g(x) &= f(g(x+k)) - f(g(x)) \\ &= f(g(x) + G_*(k)) - f(g(x))\,, \\ &= F_*(G_*(k)) \end{aligned}$$

et de même

$$\begin{aligned} (f+g)(x+k) - (f+g)(x) &= (f+g)(x) + F_*(k) + G_*(x) - (f+g)(x) \\ &= F_*(k) + G_*(k) \end{aligned}.$$

On déduit de **i)** et de **ii)** que $(F^{-1})_* = (F_*)^{-1}$, ce qui implique que $|\det(F_*)| = 1$. □

On définit le *degré* d'une application continue $F\ : \mathbf{T}^r \to \mathbf{T}^r$ comme l'entier $\deg(F) = \det(F_*)$. La proposition précédente nous dit que $\deg(F\circ G) = \deg(F)\deg(G)$ pour toutes applications continues $F$ et $G$. En particulier, $\deg(F) = \pm 1$ si $F$ est un homéomorphisme. Remarquons que $F\ : \mathbf{T}^r \to \mathbf{T}^r$ est un homéomorphisme si et seulement si $F$ est relevé par un homéomorphisme de $\mathbf{R}^r$ et $\deg(F) = \pm 1$.

Nous dirons que deux transformations continues $F\ : \mathbf{T}^r \to \mathbf{T}^r$ et $G\ : \mathbf{T}^r \to \mathbf{T}^r$ sont *homotopes* s'il existe une application continue $\Phi\ : [0,1]\times\mathbf{T}^r \to \mathbf{T}^r$ telle que pour tout $x \in \mathbf{T}^r$, on a

$$\Phi(0,x) = F(x),\ \ \Phi(1,x) = G(x).$$

Proposition 5.4.3 : *Les transformations $F$ et $G$ sont homotopes si et seulement si $F^* = G^*$.*

*Preuve.* Supposons d'abord que $F_* = G_*$. Fixons un relèvement $f$ de $F$ et un relèvement $g$ de $G$ et définissons une application continue

$$\begin{aligned} \varphi\ : [0,1]\times\mathbf{R}^r &\to \mathbf{R}^r \\ (t,x) &\mapsto tg(x) + (1-t)f(x) \end{aligned}.$$

Pour tout $t \in [0,1]$, pour tout $x \in \mathbf{R}^r$ et pour tout $k \in \mathbf{Z}^r$, on a

$$\begin{aligned} \varphi(t,x+k) &= tg(x+k) + (1-t)f(x+k) \\ &= tg(x) + tG_*(k) + (1-t)f(x) + (1-t)F_*(k)\,. \\ &= \varphi(t,x) + F_*(k) \end{aligned}$$

On en déduit que chaque application $f_t\ : x \mapsto \phi(t,x)$ relève une transformation $F_t\ : \mathbf{T}^r \to \mathbf{T}^r$ telle que $(F_t)_* = F_*$. Remarquons que $F_0 = F$ et $F_1 = G$. Il reste à prouver que

$$\begin{aligned} \Phi\ : [0,1]\times\mathbf{T}^r &\to \mathbf{T}^r \\ (t,\widehat{x}) &\mapsto F_t(\widehat{x}) \end{aligned}$$

est continue en tout point $(t_0,\widehat{x}_0)$. Fixons une norme $\|\ \|$sur $\mathbf{R}^r$ et considérons la distance associée $d$ sur $\mathbf{T}^r$. Rappelons que $\varepsilon$ a été défini dans la proposition 2.1.4. Fixons $x_0 \in \pi^{-1}(\{\widehat{x}_0\})$ et notons que $\Phi$ est continue sur $[0,1]\times B(\widehat{x}_0,\varepsilon)$ car on peut l'écrire

$$(t,\widehat{x}) \mapsto \pi(\varphi(t,{\pi_{B(x_0,\varepsilon)}}^{-1}(\widehat{x}))).$$

Supposons maintenant que $F$ et $G$ sont homotopes et considérons une application continue $\Phi\ : [0,1] \times \mathbf{T}^r \to \mathbf{T}^r$ telle que $\Phi(0,x) = F(x)$ et $\Phi(1,x) = G(x)$ pour tout $x \in \mathbf{T}^r$. Pour toute application linéaire $A\ : \mathbf{R}^r \to \mathbf{R}^r$ dont la matrice dans la base canonique est à coefficients entiers, notons $I_A$ l'ensemble des réels $t \in [0,1]$ tels que $F_t\ : \widehat{x} \mapsto \Phi(t,\widehat{x})$ vérifie $(F_t)_* = A$. On va montrer que $I_A$ est une partie ouverte de $[0,1]$. Comme les ensembles $I_A$ définissent une partition de $[0,1]$, ces ensembles seront également fermés, et puisque $[0,1]$ est connexe on en déduira que $[0,1]$ est égal à un unique ensemble $I_A$. On aura donc $(F_0)_* = (F_1)_* = A$.

Fixons une norme $\|\ \|$ sur $\mathbf{R}^r$ et considérons la distance associée $d$ sur $\mathbf{T}^r$ ainsi que le réel $\varepsilon$ défini dans la proposition 5.1.4. Fixons une application linéaire $A\ : \mathbf{R}^r \to \mathbf{R}^r$ dont la matrice dans la base canonique est à coefficients entiers et supposons que $t_0 \in I_A$. L'ensemble $[0,1] \times \mathbf{T}^r$ étant compact, $\Phi$ est uniformément continue. Ainsi, il existe $\eta > 0$ tel que $d(\Phi(t,\widehat{x}), \Phi(t',\widehat{x}')) < \varepsilon$ si $|t - t'| < \eta$ et $d(\widehat{x},\widehat{x}') < \eta$. En particulier, pour tout $t \in [0,1]$ tel que $|t - t_0| < \eta$ et pour tout $\widehat{x} \in \mathbf{T}^r$ on a $d(F_t(\widehat{x}), F_{t_0}(\widehat{x})) < \varepsilon$. On va montrer que $(F_t)_* = A$ pour ces valeurs de $t$. Posons $H = F_t - F_{t_0}$ et considérons un relèvement $h$ de $H$. Puisque, pour tout $\widehat{x} \in \mathbf{T}^r$ on a $d(H(\widehat{x}), \widehat{0}) < \varepsilon)$, on en déduit que l'image de $h$ est incluse dans $\pi^{-1}(B(\widehat{0},\varepsilon)) = \bigsqcup_{k\in\mathbf{Z}^r} B(k,\varepsilon)$. Puisque cette image est connexe, elle est incluse dans une unique boule $B(k,\varepsilon)$. La fonction $h$ est donc bornée et on a $H_* = (F_t)_* - (F_{t_0})_* = 0$. □

On en déduit immédiatement :

Corollaire 5.4.4 : *Toute application continue $F\ : \mathbf{T}^n \to \mathbf{T}^n$ est homotope à l'endomorphisme linéaire $\widehat{F}^*$.*