Aspects quantitatifs du théorème de récurrence de Poincaré

Thierry Monteil

juin 2002

Table des matières

1 Introduction

Le théorème de récurrence de Poincaré nous dit que si T est une application mesurable d'un espace mesuré X de mesure finie dans lui-même qui préserve la mesure, si A est une partie mesurable de X de mesure strictement positive, alors l'orbite de presque tout point de A repassera une infinité de fois dans A.

Une idée importante de la démonstration consiste à dire que si U est une partie de X de mesure strictement positive, les préimages itérées de U par T ne peuvent être toutes disjointes (car elles sont toutes de même mesure strictement positive et que X est de mesure finie). Donc, si x est dans $T^{-k}(U) \cap T^{-l}(U)$ avec $1 \leq k < l$, alors $T^k(x) \in U \cap T^{l-k}(U)$. (voir [Po])

Dans son ouvrage de vulgarisation *Hasard et chaos* [Ru], David Ruelle donne l'interprétation suivante de l'irréversibilité en thermodynamique :

«Vous avez soigneusement versé une couche d'eau chaude au-dessus d'une couche d'eau froide. Et vous avez obtenu de cette manière que votre système soit dans une petite région très particulière de l'espace des phases. Mais la chaleur diffuse et les couches se mélangent, et au bout d'un certain temps vous avez de l'eau uniformément tiède, correspondant à une beaucoup plus grande région de l'espace des phases. Si vous attendez assez longtemps, l'éternel retour ramènera votre système à une couche d'eau froide et une couche d'eau chaude comme au début de l'expérience. Mais combien de temps devrez-vous attendre ? Le calcul de ce temps est lié aux calculs de nombre d'états fait au chapitre 16, et la réponse est d'une énormité effrayante, atroce. Le temps qu'il faut attendre est simplement trop long. La vie est bien trop brève pour que l'on ne revoie jamais une couche d'eau chaude surmontant une couche d'eau froide, et dans ce sens, le mélange des deux couches est irréversible.
L'explication de l'irréversibilité que nous avons obtenue, en suivant Boltzmann, est à la fois simple et assez subtile. C'est une explication probabiliste. Il n'y a pas d'irréversibilité dans les lois fondamentales de la physique, mais l'état initial que nous avons choisi pour notre système a une caractéristique importante : cet état initial est *très improbable*. Je veux dire par là que son volume relatif dans l'espace des phases est très petit (ou que son entropie est petite). L'évolution temporelle conduit alors à une région de volume relativement grand (ou de grande entropie) qui correspond à un état très probable du système. En principe, après un

temps très long, le système retournera à son improbable état initial... mais nous ne verrons jamais la chose se produire. Un physicien aimera idéaliser cette situation en faisant tendre le nombre de particules du système vers l'infini, de sorte que le temps de l'éternel retour tende aussi vers l'infini. A la limite, nous avons donc une évolution temporelle vraiment irréversible.»

Le très bel article d'Etienne Ghys [Gh] nous donne aussi une évaluation (plus précise) du temps de retour lié cette fois-ci à la discrétisation de l'action d'un automorphisme hyperbolique du tore $\mathbb{T}^2$.

Ces deux textes vont dans le même sens : ils ont une estimation "pessimiste" du temps de retour de Poincaré : leur but est de nous montrer que le théorème de Poincaré est de nature qualitative et que, de façon générale, le retour est plutôt long.

L'objet de ce mémoire est l'étude de l'article de Michael D. Boshernitzan, *Quantitative recurrence results* [Bo2], qui étudie la vitesse (mesurée d'un point de vue métrique) du retour de Poincaré.

2 Un exemple

2.1 Isométries de $\mathbb{T}$

Soit T une isométrie de $\mathbb{T} = \mathbb{R}/\mathbb{Z}$.
Si T est d'ordre fini, la vitesse de retour à extraction près est infinie puisque l'on peut trouver une extraction $k \in \uparrow(\mathbb{N}, \mathbb{N})$ telle que pour tout x dans $\mathbb{T}$, $d(T^{k_n}(x), x) = 0$ $(n \in \mathbb{N})$.
Sinon, $T = R_\varphi$ est une rotation d'angle φ irrationel.
Si $x \in \mathbb{T}$ et $k \geq 1$, $d(x, T^k(x)) = d(x, x + k\varphi) = d(0, k\varphi) = ||k\varphi||$ est indépendante de $x \in \mathbb{T}$ (si $u \in \mathbb{R}$, $||u||$ désigne la distance de u à $\mathbb{Z}$).

On vient donc de se ramener à un problème d'approximation diophantienne : la théorie des fractions continues nous dit qu'il existe une suite strictement croissante d'entiers $(q_n)_{n\in\mathbb{N}}$ telle que pour tout n dans $\mathbb{N}$, $q_n||q_n\varphi|| \leq 1/\sqrt{5}$ (parmi les dénominateurs de trois convergeants successifs de φ, il y en a au moins un qui vérifie l'inégalité). Cette théorie nous dit aussi que si φ a un développement qui stationne à 1, et si $c < 1/\sqrt{5}$, il n'existe qu'un nombre fini d'entiers $q \geq 1$ tels que $q||q\varphi|| \leq c$ (voir [Ca]).

En particulier, $\liminf\limits_{n\to\infty} nd(T^n(x), x) \leq 1/\sqrt{5}$ avec égalité par exemple si $\varphi = (1 + \sqrt{5})/2$.

2.2 En dimension α

On se place maintenant dans $\mathbb{T}^\alpha$ $(\alpha \in \mathbb{N}^*)$ que l'on munit de la métrique définie par $d((x_1, \ldots, x_\alpha), (y_1, \ldots, y_\alpha)) \overset{\text{def}}{=} \max_{i=1}^\alpha d_{\mathbb{T}}(x_i, y_i)$.
Soit $\varphi = (\varphi_1, \ldots, \varphi_\alpha) \in \mathbb{R}^\alpha \setminus \mathbb{Q}^\alpha$. On considère l'application $T = \left(\begin{array}{ccc} \mathbb{T}^\alpha & \longrightarrow & \mathbb{T}^\alpha \\ x & \longmapsto & x + \varphi \end{array} \right)$.
On a encore, pour x dans $\mathbb{T}^\alpha$ et $k \geq 1$, $d(T^k(x), x) = \max_{i=1}^\alpha ||k\varphi_i||$.

Le raisonnement qui suit est emprunté à Dirichlet (voir [Di]) :
Soit n dans $\mathbb{N}^*$, on découpe $\mathbb{T}^\alpha \sim [0, 1[^\alpha$ en n^α carreaux de côtés $1/n$.
Comme $I = \{0, \ldots, n^\alpha\}$ est de cardinal $n^\alpha + 1 > n^\alpha$, par le principe des tiroirs, il existe $k < l$ dans I tels que $k\varphi$ et $l\varphi$ soient dans la même boîte et en particulier, $d(k\varphi, l\varphi) \leq 1/n$ donc $d(0, (l - k)\varphi) \leq 1/n$: on vient de trouver $k_n \overset{\text{def}}{=} l - k \leq n^\alpha$ tel que $d(T^{k_n}(x), x) \leq 1/n \leq 1/{k_n}^{1/\alpha}$. Comme $\varphi \notin \mathbb{Q}^\alpha$ et que $d(T^{k_n}(x), x) \xrightarrow[n\to\infty]{} 0$, la suite $(k_n)_{n\in\mathbb{N}^*}$ ne saurait être bornée, donc $\liminf\limits_{n\to\infty} n^{1/\alpha} d(T^n(x), x) \leq 1$.

On peut voir le raisonnement fait ici comme une version discrète de celui de Poincaré.

Réciproquement, si φ est tel que $(1, \varphi_1, \ldots, \varphi_\alpha)$ est une base d'un corps de nombres totalement réel, on sait qu'il est mal approximable au sens de Schmidt (voir [Sc]) : il existe $c > 0$ tel que pour tout $q \geq 1$, $d(0, q\varphi) \geq c/q^{1/\alpha}$, et dans ce cas, si $x \in \mathbb{T}^\alpha$, $\liminf_{n\to\infty} n^{1/\alpha} d(T^n(x), x) \geq c > 0$.

On se doute de quelque chose...

3 Théorème

3.1 Énoncé

On se donne $(X, \mathcal{A}, \mu)$ un espace mesuré de mesure totale 1.
On munit X d'une distance d telle que les ouverts de X soient dans $\mathcal{A}$.
Si $\alpha > 0$, on note H_α la α-mesure de Hausdorff relative à d sur X (voir Annexe).
Soit T une application mesurable de X dans X qui préserve la mesure μ.

Théorème 1
On suppose que, pour un certain $\alpha > 0$, $H_\alpha(X)$ est finie.
Alors, pour μ-presque tout x dans X, $\liminf_{n\to\infty} n^{1/\alpha} d(x, T^n(x)) < \infty$.

Nous démontrerons une version plus générale de ce théorème dans le paragraphe suivant.

3.2 Remarque

On a ici une propriété qui mesure la vitesse de convergence (à extraction près) de la suite des itérés de x vers x. Une autre façon de procéder (plus naturelle à priori, si l'on regarde les textes cités en introduction) est de se fixer un certain rayon $r > 0$ et de regarder le temps que met l'orbite de x à revenir dans la boule $B(x, r]$ (premier retour).
Ces deux points de vue sont en fait liés.
De façon précise, notons, pour $r > 0$ et x dans X,

$$\tau_r(x) \stackrel{\text{def}}{=} \inf\{k \in \mathbb{N}^* \mid T^k(x) \in B(x, r]\}$$

et

$$R(x) = \liminf_{r\to 0} -\frac{\log(\tau_r(x))}{\log(r)}$$

(de sorte que si $\tau_r(x) = \lfloor 1/r^7 \rfloor$, $R(x) = 7$).

Proposition 1 ([BS]) *Soient x dans X et $\alpha > 0$, on l'équivalence :*

$$(R(x) \leq \alpha) \iff (\forall \varepsilon > 0)(\liminf_{n \to \infty} n^{1/(\alpha+\varepsilon)} d(x, T^n(x)) = 0).$$

Preuve :

$\Rightarrow$: Soit $\varepsilon > 0$. On dispose donc d'une suite $(r_n)_{n \in \mathbb{N}}$ de réels strictement positifs qui tend vers 0 telle que $-\frac{\log(\tau_{r_n}(x))}{\log(r_n)} \leq \alpha + \varepsilon$ pour tout n dans $\mathbb{N}$.

En particulier, $\gamma_n \stackrel{\text{def}}{=} \tau_{r_n}(x) \leq {r_n}^{-(d+\varepsilon)}$.

Si (γ_n) est bornée, x est T-périodique et on a le résultat.

Sinon, quitte à rééxtraire, (γ_n) est strictement croissante.

On a ${\gamma_n}^{1/(\alpha+2\varepsilon)} d(x, T^{\gamma_n}(x)) \leq {r_n}^{-(\alpha+\varepsilon)/(\alpha+2\varepsilon)} r_n = r_n^{\varepsilon/(\alpha+2\varepsilon)} \xrightarrow[n \to \infty]{} 0$

2ε étant aussi arbitraire qu'ε, on a le résultat.

$\Leftarrow$: Si $n \geq 1$, on pose $r_n \stackrel{\text{def}}{=} d(x, T^n(x))$. Ainsi, $\tau_{r_n}(x) \leq n$ pour tout $n \geq 1$.

Soit $\varepsilon > 0$. On a une extraction γ telle que $\tau_{r_{\gamma_n}}(x)^{1/(\alpha+\varepsilon)} r_{\gamma_n} \xrightarrow[n \to \infty]{} 0$.

En particulier, quitte a rééxtraire, $\tau_{r_{\gamma_n}}(x)^{1/(\alpha+\varepsilon)} r_{\gamma_n} \leq 1$ pour tout n dans $\mathbb{N}$.

Ainsi, pour n dans $\mathbb{N}$, $-\dfrac{\log(\tau_{r_{\gamma_n}}(x))}{\log(r_{\gamma_n})} \leq \alpha + \varepsilon$.

En passant à la limite inférieure puis en faisant tendre ε vers 0, on a le résultat.

Dans la preuve du théorème 1, nous estimerons des temps de premier retour.

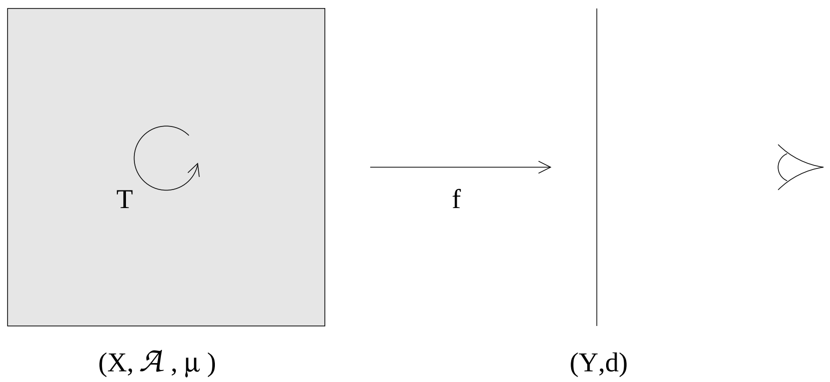
4 Raffinements

4.1 Version “projetée”

On se place maintenant dans le cadre suivant :

- $(X, \mathcal{A}, \mu)$ est un espace mesuré de mesure totale 1.
- T est une application mesurable de X dans X qui préserve la mesure μ .
- (Y, d) est un espace métrique (si $\alpha > 0$, on note encore H_α la mesure de Hausdorff associée).
- f est une application mesurable de X dans Y (muni de la σ-algèbre borélienne).

La situation peut donc se résumer ainsi :



Et on peut comprendre f comme un observable.

Par ailleurs, mettre métrique et mesure dans deux mondes différents permet d'éclaircir la démonstration et, en anticipant sur le paragraphe 6, nous fera comprendre en quoi faire coïncider μ et H_α est une hypothèse forte.

Théorème 2

On suppose que, pour un certain $\alpha > 0$, $H_\alpha(Y)$ est finie.

Alors, pour μ-presque tout x dans X, $\liminf_{n\to\infty} n^{1/\alpha} d(f(x), f(T^n(x))) < \infty$.

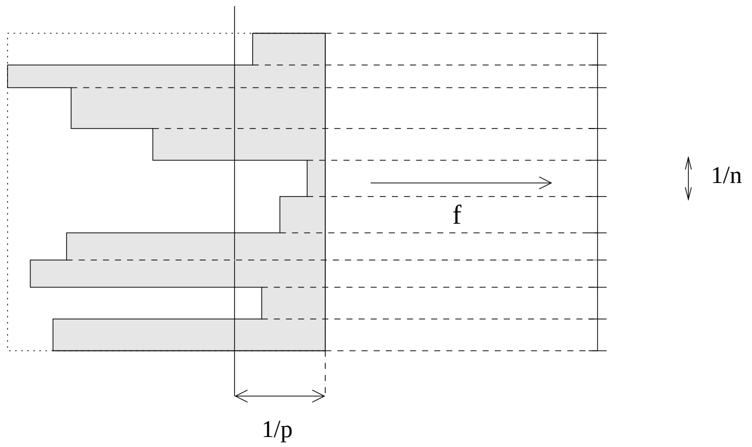
On remarque qu'en posant $Y = X$ et $f = Id_X$, on retrouve le théorème 1.

Preuve :

Soit $C > H_\alpha(Y)$. Soient $p \in \mathbb{N}^*$ et $n \in \mathbb{N}^*$.

On dispose d'un recouvrement $(A_i)_{i\in\mathbb{N}}$ de Y par des parties A_i de diamètre r_i inférieur à $1/n$ telles que $\sum_{i\in\mathbb{N}} r_i^\alpha \leq C$. Quitte à remplacer A_i par $\overline{A_i} \setminus \overline{\bigcup_{j<i} A_j}$ ($i \in \mathbb{N}$), on peut supposer les A_i disjoints et mesurables.

Soit $I \stackrel{\text{def}}{=} \{i \in \mathbb{N} \mid \mu(f^{-1}(A_i)) \leq \frac{1}{p}\text{diam}(A_i)^\alpha = \frac{1}{p}{r_i}^\alpha\}$ et soit $J \stackrel{\text{def}}{=} \mathbb{N} \setminus I$.



Si $j \in J$, on note $C_j = \{x \in f^{-1}(A_j) \mid (\exists k \in \{1, \ldots, \left[\frac{p^2}{{r_j}^\alpha}\right]\})(T^k(x) \in f^{-1}(A_j))\}$ l'ensemble des éléments de $f^{-1}(A_j)$ dont on contrôle le temps de retour (relativement au taux de remplissage p et à la précision n). Soit $D_j \stackrel{\text{def}}{=} f^{-1}(A_j) \setminus C_j$.

Montrons que $\mu(D_j) \leq \frac{1}{p}\mu(f^{-1}(A_j))$ (pour j dans J).

J'affirme que les $T^{-k}(D_j)$ sont disjoints pour $0 \leq k \leq \left[\frac{p^2}{{r_j}^\alpha}\right] \stackrel{\text{not}}{=} K_j$:

> Dans le cas contraire, on a x dans $T^{-k}(D_j) \cap T^{-l}(D_j)$ avec $0 \leq k < l \leq K_j$ donc $T^k(x) \in D_j \cap T^{l-k}(D_j)$ avec $l - k \in \{1, \ldots, K_j\}$, ce qui est contraire la définition de D_j.

Ainsi,

$$1 = \mu(X) \geq \mu\left(\bigcup_{k=0}^{K_j} T^{-k}(D_j)\right) = \sum_{k=0}^{K_j} \mu(T^{-k}(D_j)) = (K_j+1)\mu(D_j) \geq \frac{p^2}{{r_j}^{\alpha}}\mu(D_j)$$

donc $\mu(D_j) \leq \dfrac{1}{p}\dfrac{{r_j}^{\alpha}}{p} \leq \dfrac{1}{p}\mu(f^{-1}(A_j))$.
Nous pouvons remarquer que le raisonnement fait ici est exactement le même que celui fait dans la démonstration du théorème de Poincaré.

Si x est dans C_j $(j \in J)$, il existe k dans $\{1, \ldots, \left[\frac{p^2}{{r_j}^{\alpha}}\right]\}$ tel que $d(f(T^k(x)), f(x)) \leq r_i \leq \frac{1}{n}$. Pour un tel k, on a en particulier $r_i \leq \frac{p^{2/\alpha}}{k^{1/\alpha}}$.
Posons $C_{p,n} \stackrel{\text{def}}{=} \bigcup_{j\in J} C_j$ (les C_j dépendent de p et n).
Si $x \in C_{p,n}$, on a donc $k \geq 1$ tel que $d(f(T^k(x)), f(x)) \leq \min(\frac{p^{2/\alpha}}{k^{1/\alpha}}, \frac{1}{n})$.
$C_{p,n}$ est de mesure supérieure à $1 - \sum_{i\in I} \mu(f^{-1}(A_i)) - \sum_{j\in J} \frac{1}{p}\mu(f^{-1}(A_j)) \geq 1 - \frac{C}{p} - \frac{1}{p} = 1 - \frac{C+1}{p}$.

L'ensemble $C_{p,\infty}$ des x de X pour lesquels il existe une infinité de n dans $\mathbb{N}^*$ tels que $x \in C_{p,n}$ est $\bigcap_{N\in\mathbb{N}^*} \bigcup_{n\geq N} C_{p,n}$ qui est de mesure supérieure à $1 - \frac{C+1}{p}$.
Soit enfin $C_{\cup,\infty} \stackrel{\text{def}}{=} \bigcup_{p\geq 1} C_{p,\infty}$.

discussion

Supposons f injective. Soit x dans $C_{\cup,\infty}$.

Il y a deux cas à considérer :

1. S'il existe k dans $\mathbb{N}^*$ tels que $f(T^k(x)) = f(x)$, alors x est T-périodique et à fortiori, $\liminf\limits_{k\to\infty} k^{1/\alpha} d(f(x), f(T^k(x))) = 0 < \infty$.
2. Sinon, par définition de $C_{\cup,\infty}$, il existe une constante L ($L = p^{2/\alpha}$ où p est tel que $x \in C_{p,\infty}$) et une extraction $\gamma \in \uparrow(\mathbb{N}, \mathbb{N}^*)$ telle que pour tout n dans $\mathbb{N}$, il existe $k = k_n \geq 1$ tel que $d(f(T^k(x)), f(x)) \leq \min(\frac{L}{k^{1/\alpha}}, \frac{1}{\gamma_n})$.
   La suite (k_n) tend vers l'infini car pour tout $k \geq 1$, $f(T^k(x)) \neq f(x)$ et $d(f(T^{k_n}(x)), f(x)) \xrightarrow[n\to\infty]{} 0$; quitte à réextraire, on peut la supposer strictement croissante, et on vient de trouver une extraction $(k_n)_{n\in\mathbb{N}}$ telle que pour tout n dans $\mathbb{N}$, $k_n^{1/\alpha} d(f(T^{k_n}(x)), f(x)) \leq L$.

Dans le cas général, l'alternative précédente n'a pas forcément lieu : il se peut qu'il existe un $k \geq 1$ tel que $f(T^k(x)) = f(x)$ sans qu'il en existe une infinité et cela peut même se produire sur un ensemble de mesure strictement positive.

En effet, si on considère, pour $\varphi \in \mathbb{R} \setminus \mathbb{Q}$,
$T_\varphi \stackrel{\text{def}}{=} \begin{pmatrix} \mathbb{T} \times \{0,1\} & \longrightarrow & \mathbb{T} \times \{0,1\} \\ (\theta, 0) & \longmapsto & (\theta, 1) \\ (\theta, 1) & \longmapsto & (\theta + \varphi, 0) \end{pmatrix}$ et $f \stackrel{\text{def}}{=} \begin{pmatrix} \mathbb{T} \times \{0,1\} & \longrightarrow & \mathbb{T} \\ (\theta, \varepsilon) & \longmapsto & \theta \end{pmatrix}$, tout point x de la forme $(\theta, 0)$ vérifie $f(T_\varphi(x)) = f(x)$ et $(\forall k \geq 2)(f(T_\varphi^k(x)) \neq f(x))$.

Je propose la solution suivante :

Si T' est une application qui préserve la mesure, on a vu qu'il existait un ensemble $C_{\cup,\infty}(T')$ de mesure 1 tel que pour tout x dans $C_{\cup,\infty}(T')$, il existe une constante L et une extraction $\gamma \in \uparrow(\mathbb{N}, \mathbb{N}^*)$ (qui dépendent de x et de T') telles que pour tout n dans $\mathbb{N}$, il existe $k_n \geq 1$ tel que $d(f(T^{k_n}(x)), f(x)) \leq \min(\frac{L}{{k_n}^{1/\alpha}}, \frac{1}{\gamma_n})$.

Soit $\mathcal{C} \stackrel{\text{def}}{=} \bigcap_{K \geq 1} C_{\cup,\infty}(T^K)$ (les T^K préservent la mesure). $\mathcal{C}$ est de mesure 1.

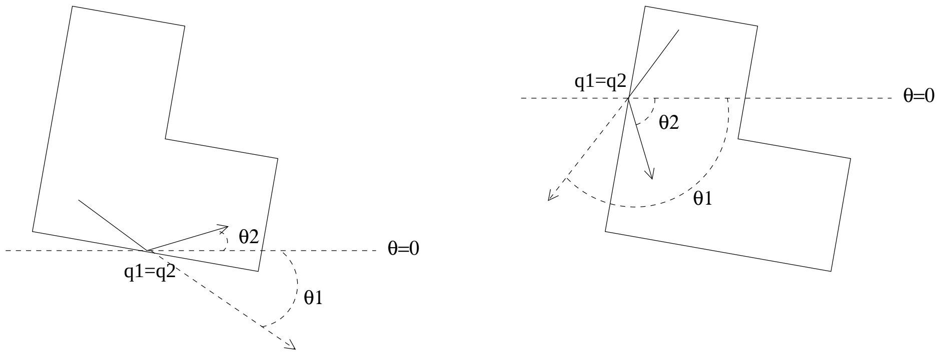
Si x est dans $\mathcal{C}$, on se retrouve face à deux cas :

- il existe une infinité d'entiers $k \geq 1$ tels que $f(T^k(x)) = f(x)$ et on applique le raisonnement fait au cas 1 plus haut.
- il existe $K \geq 1$ tel que $(\forall k \geq K)(f(T^k(x)) \neq f(x))$ et dans ce cas, on peut appliquer le second cas plus haut à $T' = T^K$ de sorte qu'il existe une constante L et une extraction $(k_n)_{n \in \mathbb{N}}$ telle que pour tout n dans $\mathbb{N}$, $k_n^{1/\alpha} d(f(T^{Kk_n}(x)), f(x)) \leq L$ et en particulier $\liminf_{k \to \infty} k^{1/\alpha} d(f(T^k(x)), f(x)) \leq \dfrac{L}{K^{1/\alpha}} < \infty$ (considérer l'extraction $(Kk_n)_{n \in \mathbb{N}}$).

Une application directe : le billard

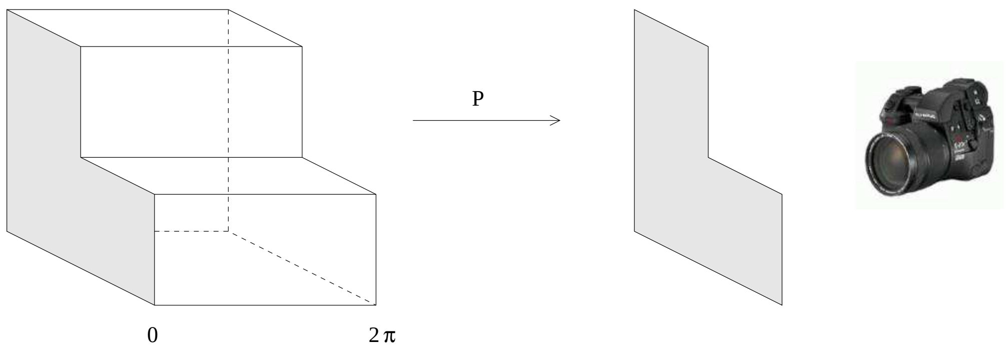
Soit maintenant Q une table de billard polygonale (en forme de L par exemple). Supposons que l'on y lance une bille, que l'on prenne une photo toutes les secondes et qu'on se demande au bout de combien de temps on verra la bille revenir près de son point de départ.

La bille (et la direction qu'elle emprunte) peut être décrite comme un point x de $X = Q \times \mathbb{S}^1$ quotienté correctement (on identifie les points (q_1, θ_1) et (q_2, θ_2) si $q_1 = q_2$ est sur le bord de Q et si les angles θ_1 et θ_2 sont tels que les lois de reflexions habituelles soient vérifiées : voir [Ta] pour une définition précise).



Son déplacement est un flot noté φ^t (qui préserve la mesure de Lebesgue sur X).

Prendre une photo toutes les secondes revient non seulement à regarder l'application $T = \varphi^1$ mais aussi à projeter X sur $Y = Q$ (en effet, la photo, pourvu que le temps d'exposition soit suffisament court (image nette), "oublie" la direction de la bille et ne retient que sa position sur la table).



On peut alors affirmer que pour presque-toute condition initiale, quitte à selectionner un album $(p_n)_{n\in\mathbb{N}^*}$ de photos, on verra la bille revenir vers son point de départ en $O(1/\sqrt{p_n})$. On remarque ici que c'est la dimension de Y qui compte et non celle de l'espace où s'effectue la transformation.

4.2 Variations sur H_α

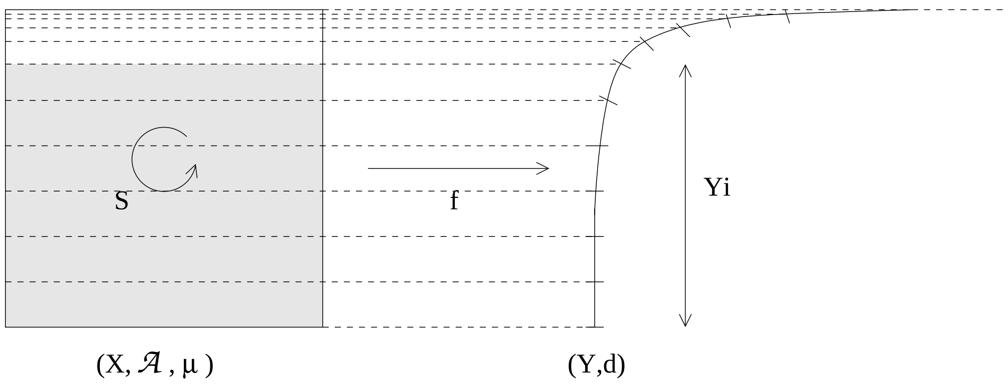
$H_\alpha(Y) = 0$

Proposition 2 *On suppose que, pour un certain $\alpha > 0$, $H_\alpha(Y) = 0$.*
Alors, pour μ-presque tout x dans X, $\liminf_{n\to\infty} n^{1/\alpha} d(f(x), f(T^n(x))) = 0$.

La preuve est presque la même que celle faite dans le paragraphe précédent.

$H_\alpha(Y)$ est σ-finie

On suppose ici que $H_\alpha(Y)$ est σ-finie : on dispose donc d'une suite $(Y_i)_{i\in\mathbb{N}}$ d'ensembles mesurables de H_α-mesure finie telle que $Y = \bigcup_{i\in\mathbb{N}} Y_i$.
Soit i dans $\mathbb{N}$ tel que $\mu(f^{-1}(Y_i)) > 0$. Le théorème de récurrence de Poincaré nous dit que pour μ-presque tout x de $f^{-1}(Y_i)$, l'ensemble $N(x) \stackrel{\text{def}}{=} \{n \geq 1 \mid T^n(x) \in f^{-1}(Y_i)\}$ est infini.
On peut donc construire une extraction $(n_k(x))_{k\in\mathbb{N}}$ d'image $N(x)$.
Soit alors $S \stackrel{\text{def}}{=} \begin{pmatrix} f^{-1}(Y_i) & \longrightarrow & f^{-1}(Y_i) \\ x & \longmapsto & T^{n_0(x)}(x) \end{pmatrix}$ l'application induite par T sur $f^{-1}(Y_i)$ qui est définie presque partout et préserve la mesure induite sur Y_i (voir [Kr]).



Le théorème 2 s'applique à S donc pour μ-presque tout x de $f^{-1}(Y_i)$,
$\liminf_{n\to\infty} n^{1/\alpha} d(f(x), f(S^n(x))) < \infty$.
Soit x un point de $f^{-1}(Y_i)$ tel que $N(x)$ soit infini.
Il en est de même pour l'orbite de x par S avec, si $k \geq 1$, $n_0(S^k(x)) = n_k(x) - n_{k-1}(x)$ donc $S^k(x) = S^{k-1}(S(x)) = S^{k-2}(S(T^{n_0(x)}(x))) = S^{k-2}(T^{n_1(x)}(x)) = \cdots = T^{n_{k-1}(x)}(x)$.
Le lemme de Kac nous dit que n_0 est une fonction L^1 sur $f^{-1}(Y_i)$ donc le théorème de Birkhoff s'applique et pour μ-presque tout x de $f^{-1}(Y_i)$,

$$\frac{1}{n} n_{n-1}(x) = \frac{1}{n} \sum_{k=0}^{n-1} n_0(S^k(x)) \xrightarrow[n\to\infty]{} g(x)$$

où g est une fonction L^1 donc finie μ-presque partout.
Ainsi, pour μ-presque tout x de $f^{-1}(Y_i)$, il existe une constante $L(x) \geq 0$ telle que pour tout $n \geq 1$, $n_{n-1}(x) \leq nL(x)$ de sorte que
$\liminf_{n\to\infty} n^{1/\alpha} d(f(x), f(T^n(x))) \leq \liminf_{n\to\infty} n_{n-1}(x)^{1/\alpha} d(f(x), f(T^{n_{n-1}(x)}(x))) \leq$
$\liminf_{n\to\infty} L(x)^{1/\alpha} n^{1/\alpha} d(f(x), f(S^n(x)))$.

On en déduit la

Proposition 3 *On suppose que, pour un certain $\alpha > 0$, $H_\alpha(Y)$ est σ-finie.*
Alors, pour μ-presque tout x dans X, $\liminf_{n\to\infty} n^{1/\alpha} d(f(x), f(T^n(x))) < \infty$.

On remarque que si $H_\alpha(Y)$ est infini et σ-finie, alors $\alpha = \dim(Y)$.
En effet, comme $H_\alpha(Y)$ est infini, $\alpha \leq \dim(Y)$.
Si $\alpha < \dim(Y)$, soit $\beta \in]\alpha, \dim(Y)[$ de sorte que $H_\beta(Y) = \infty$. Comme $H_\alpha(Y)$ est σ-finie, il existe une suite $(Y_i)_{i\in\mathbb{N}}$ d'ensembles mesurables de H_α-mesure finie telle que $Y = \bigcup_{i\in\mathbb{N}} Y_i$.
Or comme $H_\alpha(Y_i)$ est finie, $H_\beta(Y_i) = 0$ $(i \in \mathbb{N})$ donc $H_\beta(Y) = 0$ ce qui n'est pas.

5 Échanges d'intervalles

En contraposant le théorème 1, on voit que le contrôle de la vitesse de convergence permet d'estimer certaines dimensions.

5.1 Préliminaire diophantien

Soit $U = \{u_1, \ldots, u_n\} \subset \mathbb{R}^*$ un ensemble de cardinal $n \geq 1$; on pose, pour $N \in \mathbb{N}^*$,

$$\Psi(U, N) \stackrel{\text{def}}{=} \{\sum_{i=1}^{n} k_i u_i \quad | \quad |k_i| \leq N\}$$

puis

$$\Phi(U) \stackrel{\text{def}}{=} \limsup_{N \to \infty} -\frac{\log(\min(\Psi(U, N) \cap \mathbb{R}_+^*))}{\log(N)}.$$

On cherche à estimer $\Phi(U)$ lorsque U est composé de nombres algébriques.
Familiarisons-nous avec cette notion avant d'énoncer précisement la proposition.
Voici quelques propriétés :

1. $\Psi(U, N)$ est symétrique par rapport à 0 donc stable par valeur absolue.
2. Si $\lambda \neq 0$, $\Phi(U) = \Phi(\lambda U)$, car $\min(\Psi(\lambda U, N) \cap \mathbb{R}_+^*) = |\lambda| \min(\Psi(U, N) \cap \mathbb{R}_+^*)$.
3. Si $\text{card}(U) = 1$, $\Phi(U) = 0$; on suppose désormais $n \geq 2$.

4. Si U est une partie $\mathbb{Q}$-libre, $\Phi(U) \geq n - 1$.
   En effet, les $\sum_{i=1}^{n} k_i u_i$ sont distincts et $\Psi(U, N)$ est de cardinal $(2N + 1)^n$.
   Si $M = \max_{i=1}^{n} |u_i|$, $\Psi(U, N) \subset [-nNM, nNM]$ qui est de longueur $2nNM$: on le recouvre par $(2N)^n < (2N + 1)^n$ segments de longueur $2nNM/(2N)^n$: par le principe des tiroirs de Dirichlet, il existe $a \neq b$ dans $\Psi(U, N)$ tels que $|a - b| \leq 2nNM/(2N)^n$.
   Or $|a - b| \in \Psi(U, 2N)$ donc $\min(\Psi(U, 2N) \cap \mathbb{R}_+^*) \leq 2nNM/(2N)^n$.
   Ainsi, $-\frac{\log(\min(\Psi(U,2N) \cap \mathbb{R}_+^*))}{\log(2N)} \geq \frac{-\log(2N) - \log(nM) + n\log(2N)}{\log(2N)} = n - 1 - \frac{\log(nM)}{\log(2N)}$.
   En passant à la limite supérieure, $\Phi(U) \geq n - 1$.

5. Si U et V sont finis et tels que $\text{Vect}_{\mathbb{Q}}(V) \subset \text{Vect}_{\mathbb{Q}}(U)$, alors $\Phi(V) \leq \Phi(U)$.

En effet, si $U = \{u_1, \ldots, u_n\}$ et $V = \{v_1, \ldots, v_m\}$, chaque v_i peut s'écrire $\sum_{j=1}^n \lambda_{ij} u_j$ avec $\lambda_{ij} = p_{ij}/q$ ($p_{ij} \in \mathbb{Z}$, $q \in \mathbb{N}^*$). On pose alors $P \stackrel{\text{def}}{=} \max_{1 \leq i \leq n, 1 \leq j \leq m} |p_{ij}|$.
Soit $x = \sum_{i=1}^m k_i v_i$ dans $\Psi(V, N)$ ($|k_i| \leq N \in \mathbb{N}^*$).
On a $x = \frac{1}{q} \sum_{i=1}^m \sum_{j=1}^n k_i p_{ij} u_j = \frac{1}{q} \sum_{j=1}^n (\sum_{i=1}^m k_i p_{ij}) u_j$ avec $|\sum_{i=1}^m k_i p_{ij}| \leq mPN$.
Donc $x \in \frac{1}{q}\Psi(U, mPN)$ donc $\Psi(V, N) \subset \frac{1}{q}\Psi(U, mPN)$ ($N \geq 1$).
On obtient le résultat en passant au log et à la limite supérieure.

6. Si $\dim_{\mathbb{Q}}(\mathrm{Vect}_{\mathbb{Q}}(U)) = r \stackrel{\text{not}}{=} \mathrm{rg}(U) \geq 2$, alors $\Phi(U) \geq r - 1$.
   En effet, il suffit de choisir dans U une base de $\mathrm{Vect}_{\mathbb{Q}}(U)$ et d'appliquer 5 puis 4.

7. Si U est libre sur $\mathbb{Q}$ et composé uniquement d'algébriques, $\Phi(U) = n - 1$.
   En effet, l'ensemble des algébriques étant un corps, $(1, \alpha_1, \ldots, \alpha_{n-1}) \stackrel{\text{def}}{=} \frac{1}{u_1}(u_1, \ldots, u_n)$ est aussi composé d'algébriques et demeure libre.
   Soit $\varepsilon > 0$ fixé : une généralisation du théorème de Roth pour l'approximation simultanée nous dit qu'il existe $N_\varepsilon \in \mathbb{N}$ tel que pour tout $N \geq N_\varepsilon$, si $q_1, \ldots, q_{n-1}$ sont inférieurs à N et $p \in \mathbb{Z}$, alors $|q_1\alpha_1 + \cdots + q_{n-1}\alpha_{n-1} - p| \geq \frac{1}{N^{n-1+\varepsilon}}$ (voir [Ba]).
   Ainsi, $\min(\Psi(\frac{1}{u_1}U, N) \cap \mathbb{R}_+^*) \geq \frac{1}{N^{n-1+\varepsilon}}$.
   En passant au log et à la limite supérieure, on a $\Phi(U) = \Phi(\frac{1}{u_1}U) \leq n - 1 + \varepsilon$.
   ε étant quelconque, on a le résultat.

On en déduit la

Proposition 4 *Si U est un ensemble fini de nombres algébriques, alors*

$$\Phi(U) = rg(U) - 1 \leq card(U) - 1.$$

5.2 Résultat

Pour les définitions concernant les échanges d'intervalles, on consultera [FGM] ou [CFS].

Théorème 3
Soit $T : I \to I$ un échange de r intervalles sans point périodique tel que les longueurs des intervalles échangés soient algébriques.
Alors pour toute mesure de probabilité T-invariante μ, $\dim(\mu) \geq 1/(r-1)$ (voir Annexe).

Preuve :

$U \overset{\text{def}}{=} \{T(x) - x \mid x \in I\}$ est fini et inclus dans le $\mathbb{Q}$-espace vectoriel engendré par les longueurs des intervalles échangés. Ainsi, d'après la proposition 4, $\Phi(U) \leq r - 1$.

Si $x \in I$ et $n \geq 1$,
$T^n(x) - x = (T^n(x) - T^{n-1}(x)) + (T^{n-1}(x) - T^{n-2}(x)) + \cdots + (T(x) - x) \in \Psi(U, n)$, de sorte que, comme $T^n(x) \neq x$, $|T^n(x) - x| \geq \min(\Psi(U, n) \cap \mathbb{R}_+^*)$.

Soit $\varepsilon > 0$ fixé : par définition de $\Phi(U)$, il existe $N \geq 1$ tel que pour tout $n \geq N$, $\min(\Psi(U, n) \cap \mathbb{R}_+^*) \geq n^{-\Phi(U)-\varepsilon} = \frac{n^\varepsilon}{n^{\Phi(U)+2\varepsilon}}$.
Ainsi, pour x dans I, $n^{\Phi(U)+2\varepsilon}|T^n(x) - x| \xrightarrow[n\to\infty]{} \infty$.

Supposons par l'absurde que $\dim(\mu) < \frac{1}{\Phi(U)+2\varepsilon}$.
On dispose donc d'une partie A de I telle que $\mu(A) = 1$ et $\dim(A) = \dim(\mu)$.
Soit $\alpha \in]\dim(\mu), \frac{1}{\Phi(U)+2\varepsilon}[$. Alors, $H_\alpha(A) = 0$.
Soit $X \overset{\text{def}}{=} \bigcap_{n\in\mathbb{N}} T^{-n}(A) \subset A$; X est stable par T et $\mu(X) = 1$ (car T préserve μ).
Le théorème 1 s'applique à $T|_X$ de sorte que pour μ-presque tout x dans X, $\infty = \lim\limits_{n\to\infty} n^{\Phi(U)+2\varepsilon}|T^n(x) - x| \leq \liminf\limits_{n\to\infty} n^{1/\alpha}|T^n(x) - x| < \infty$. Contradiction.

Ainsi, $\dim(\mu) \geq \dfrac{1}{\Phi(U) + 2\varepsilon}$ et, ε étant quelconque, $\dim(\mu) \geq \dfrac{1}{\Phi(U)} \geq \dfrac{1}{r-1}$.

6 Lorsque mesure et métrique coïncident

6.1 Énoncé

Le cadre est celui du paragraphe 3.

Théorème 4

On suppose que, pour un certain $\alpha > 0$, $\mu = H_\alpha$ sur les boréliens de (X, d).
Alors, pour μ-presque tout x dans X, $\liminf_{n\to\infty} n^{1/\alpha} d(x, T^n(x)) \leq 1$.

6.2 Optimalité ?

Par rapport au théorème 1, le théorème 4 nous donne une majoration explicite et indépendante du point x. On peut se demander dans quelle mesure la constante 1 est optimale.

Soit $X \stackrel{\text{def}}{=} \{0,1\}^{\mathbb{N}}$ et

$$T \stackrel{\text{def}}{=} \left(\begin{array}{ccc} X & \longrightarrow & X \\ (1,1,1,1,\dots) & \longmapsto & (0,0,0,0,\dots) \\ (\underbrace{1,1,\dots,1}_{i}, 0, \varepsilon_{i+1}, \varepsilon_{i+2}, \dots) & \longmapsto & (\underbrace{0,0,\dots,0}_{i}, 1, \varepsilon_{i+1}, \varepsilon_{i+2}, \dots) \quad (i \in \mathbb{N}) \end{array} \right)$$

la “machine dyadique” que l'on peut voir comme une incrémentation généralisée (le bit de poids le plus faible étant à gauche) : si $x \in X$ stationne à 0, on peut voir x comme le développement d'un entier n en binaire et alors $T(x)$ est le développement de $n+1$.

On munit X de la distance $d(x,y) \stackrel{\text{def}}{=} 0$ si $x = y$ et $d(x,y) \stackrel{\text{def}}{=} 2^{-\max\{k \geq 0 \mid (\forall i \leq k-1)(x_i = y_i)\}}$ sinon ; on retrouve alors la topologie produit.

Si $(x_0, \dots, x_{n-1})$ est dans $\{0,1\}^n$ $(n \geq 1)$, il définit le cylindre de longueur n

$$[x_0, \dots, x_{n-1}] \stackrel{\text{def}}{=} \{y \in X \mid (\forall i \leq n-1)(x_i = y_i)\}.$$

On définit une mesure sur les boréliens de X par $\mu([x_0, \dots, x_{n-1}]) \stackrel{\text{def}}{=} 2^{-n}$.
Pour montrer que $\mu = H_1(d)$, il suffit de le vérifier sur les cylindres en remarquant que pour toute partie A_i de X, si C_i est le plus petit (i.e. le plus long) cylindre qui contient A_i, alors $\operatorname{diam}(A_i) = \operatorname{diam}(C_i)$ de sorte que l'on peut ne considérer que des recouvrements par des cylindres.

T préserve μ car

– $T^{-1}([\underbrace{0,0,\ldots,0}_{i}]) = [\underbrace{1,1,\ldots,1}_{i}]$

– $T^{-1}([\underbrace{0,0,\ldots,0}_{i},1,\varepsilon_{i+1},\varepsilon_{i+2},\ldots,\varepsilon_{n-1}]) = [\underbrace{1,1,\ldots,1}_{i},0,\varepsilon_{i+1},\varepsilon_{i+2},\ldots,\varepsilon_{n-1}]$ $(i \in N)$

On constate que T et T^{-1} induisent une permutation circulaire sur les 2^N cylindres de longueur N pour tout $N \geq 1$. Ainsi, pour tout $x \in X$ et tout $k \in \{2^{N-1},\ldots,2^N-1\}$, $d(x,T^k(x)) \geq 2^{-N+1} \geq 1/k$, de sorte que $\liminf_{n\to\infty} nd(x,T^n(x)) \geq 1$.

Par ailleurs, cette constatation rend évident le fait que μ est la seule mesure borélienne invariante par T (on dit alors que (X,T) est uniquement ergodique).

On vient de montrer l'optimalité du théorème pour $\alpha = 1$. Si α est strictement positif, il suffit de remplacer d par $d_\alpha \stackrel{\text{def}}{=} d^{1/\alpha}$ et alors $\mu = H_\alpha(d_\alpha)$.

Ainsi, pour tout $\alpha > 0$, on peut trouver un espace et une application qui ne permettent pas d'améliorer la constante 1 dans le théorème.
L'espace qu'on a considéré n'a que peu de géométrie et si, dans le cas général, on peut dire que le théorème est optimal, on n'en sait actuellement rien pour $X = \mathbb{T}$ muni de la mesure de Lebesgue (qui est égale à H_1) sinon que la constante optimale (valable pour toutes les applications mesurables de $\mathbb{T}$ qui préservent la mesure) se trouve quelque part entre $1/\sqrt{5}$ et 1 comme on l'a vu au paragraphe 2.

6.3 Dynamique symbolique

X est encore $\{0,1\}^{\mathbb{N}}$, mais $T \stackrel{\text{def}}{=} \begin{pmatrix} X & \longrightarrow & X \\ x = (x_0,x_1,\ldots,x_n,\ldots) & \longmapsto & (x_1,x_2,\ldots,x_{n+1},\ldots) \end{pmatrix}$
désigne le shift.
Si $n \geq 1$, P_n désigne la projection $\begin{pmatrix} X & \longrightarrow & \{0,1\}^n \\ x = (x_0,x_1,\ldots,x_l,\ldots) & \longmapsto & (x_0,x_1,\ldots,x_{n-1}) \end{pmatrix}$.

On étudie le comportement de

$$R_n(x) \stackrel{\text{def}}{=} \min\{k \geq 1 \mid P_n(T^k(x)) = P_n(x)\},$$

le temps que met x pour retourner dans le cylindre $[P_n(x)] = B(x,2^{-n}]$ $(x \in X$ et $n \geq 1)$.

Théorème 5

Soit μ une mesure T-invariante qui

1. *charge les ouverts,*
2. *est sans atome.*

Alors pour μ-presque tout x dans X, $\liminf_{n\to\infty} \mu([P_n(x)])R_n(x) \leq 1$.

Preuve :

On introduit la métrique suivante sur X : $d(x,y) \stackrel{\text{def}}{=} 0$ si $x = y$ et $d(x,y) \stackrel{\text{def}}{=} \mu([P_n(x)])$ si $n = \max\{k \geq 0 \mid (\forall i \leq k-1)(x_i = y_i)\}$ sinon ; la séparation est donnée par la première hypothèse faite sur μ et on retrouve alors la topologie produit.

De même qu'au paragraphe 6.2, on montre que $\mu = H_1(d)$ (la deuxième hypothèse faite sur μ nous dit que si $(C_n)_{n\in\mathbb{N}}$ est une suite srictement décroissante de cylindres, alors $\mu(C_n) \xrightarrow[n\to\infty]{} 0$ et on constate que pour un cylindre, diamètre=mesure).

Il nous reste à appliquer le théorème 4 et à ientifier les termes en notant que $T^k(x) \in [P_n(x)] \iff d(x, T^k(x)) \leq \text{diam}([P_n(x)]) = \mu([P_n(x)])$ et en séparant les cas selon la périodicité de x (tous les détails ont déjà été rencontrés précédemment).

De même qu'au paragraphe 5, un certain contrôle de la récurrence nous permet d'étudier l'ensemble des mesures invariantes.

Soit X une partie non vide et fermée de $\{0,1\}^{\mathbb{N}}$, stable par T qui est minimale pour ces propriétés. On note encore T pour $T|_X$.

X est dit *linéairement récurrent* s'il existe une constante $C \geq 0$ telle que pour tout x de X et tout $k \geq 1$, $R_k(x) \leq Ck$.

Proposition 5 ([Du])

Si (X,T) est linéairement récurrent, alors il est uniquement ergodique.

Preuve :

Soit μ une mesure invariante ergodique pour (X,T).

Soit $b = (b_0, \ldots, b_{l-1}) \in \{0,1\}^l$ un mot fini qui apparait comme sous-mot d'un élément de X (on note $b \sqsubset X$). Soient $n \geq 1$ et $x \in X$.

$\frac{1}{n}\sum_{i=0}^{n-1}\chi_{[b]}(S^i(x)) = \frac{1}{n}\mathrm{card}\{i \in \{0,\dots,n-1\} \mid P_l(S^i(x)) = b\} \geq \frac{1}{n}\left[\frac{n}{Cl}\right] \xrightarrow[n\to\infty]{} \frac{1}{Cl} > 0.$

Par ailleurs, le théorème de Birkhoff nous dit que pour μ-presque tout x dans X, $\frac{1}{n}\sum_{i=0}^{n-1}\chi_{[b]}(S^i(x)) \xrightarrow[n\to\infty]{} \mu([b])$, de sorte que $\mu([b]) \geq \frac{1}{Cl}$.

Ainsi, si on pose $e_n \overset{\text{def}}{=} \min\limits_{b=(b_0,\dots,b_{n-1})\sqsubset X} \mu([b])$, alors $ne_n \geq \frac{1}{C} \underset{n\to\infty}{\not\to} 0$ et d'après un théorème de Boshernitzan, (X,T) est uniquement ergodique (voir [FGM] ou [Bo1]).

6.4 Ouverture

Ce qui suit est largement admis (voir [CT], [OW]).

On modélise une source d'information comme suit :

Soit $(X, \mathcal{A}, \mu)$ un espace mesuré de mesure totale 1 et T une application mesurable de X dans X qui préserve la mesure μ et qui est ergodique.

Soit $\mathcal{P} : X \to \{0, 1\}$ une application mesurable (partition).

Tout point x de X produit un mot infini

$$w(x) \stackrel{\text{def}}{=} (\mathcal{P}(x), \mathcal{P}(T(x)), \mathcal{P}(T^2(x)), \dots) \in \{0, 1\}^{\mathbb{N}}.$$

Si $x \in X$, on veut compresser les n premières lettres de $w(x)$, à savoir $y = P_n(w(x))$.

On découpe de proche en proche y en sous-mots $y_1 \dots y_p$ de sorte que chaque y_i soit différent des précédents mais que y_i privé de sa dernière lettre soit l'un des y_j avec $j < i$.

Ainsi, au lieu de coder y_i, on code sa dernière lettre et l'adresse j.

Par exemple, si $y = 1011010100010$, il sera découpé en $1, 0, 11, 01, 010, 00, 10$ et codé en $(000, 1)(000, 0)(001, 1)(010, 1)(100, 0)(010, 0)(001, 0)$.

La décompression est facile de proche en proche.

Si $c_n(x) \stackrel{\text{def}}{=} p$ désigne le nombre de sous-mots obtenus, le fichier comportera $c_n(x)(\lceil \log_2(c_n(x)) \rceil + 1)$ bits (pour chaque y_i, il faut $\lceil \log_2(c_n(x)) \rceil$ bits pour coder j et 1 bit pour coder la dernière lettre de y_i). Le taux de compression est alors $\text{taux}(x, n) \stackrel{\text{def}}{=} c_n(x)(\lceil \log_2(c_n(x)) \rceil + 1)/n$.

Le théorème de Wyner-Ziv nous dit que pour μ-presque tout x dans X, $\text{taux}(x, n)$ converge vers l'entropie de la source d'information que l'on peut définir rapidement grâce au théorème de Shannon-McMillan-Breiman : il existe $h = h(X, \mathcal{A}, \mu, T, \mathcal{P}) \in \mathbb{R}_+$ tel que pour μ-presque tout x dans X,

$$-\frac{1}{n} \log_2(\mu(w^{-1}([P_n(w(x))]))) \xrightarrow[n \to \infty]{} h.$$

La théorie de l'information nous dit que l'on ne peut pas espérer mieux faire.

L'algorithme présenté ici est celui de Lempel-Ziv : il est utilisé dans la majorité des formats de compression non-destructive actuels (gif, zip, arj, rar...). Il réalise en pratique des taux de compression d'environ 1/2 et sert aussi pour la transmission de données.

Un théorème d'Ornstein-Weiss complète le tableau : pour μ-presque tout x dans X, pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}^*$ tel que pour tout $n \geq N$,

$$\sum_{j \in \{i \leq c_n(x) \ | \ |y_i| \geq \frac{\log_2(n)}{h+\varepsilon}\}} \geq (1-\varepsilon)n.$$

Autrement dit, avec probabilité 1 et asymptotiquement, une grande proportion du mot $y = (x_0, \ldots, x_{n-1})$ est recouverte par des mots y_i de longueur presque supérieure à $\log_2(n)/h$.

Ce théorème découle du théorème suivant qui lie temps de retour et entropie :
Si $x \in X$ et $n \geq 1$, on note

$$R(x,n) \stackrel{\text{def}}{=} \min\{k \geq 1 \mid P_n(w(T^k(x))) = P_n(w(x))\}.$$

Théorème 6 ([OW]) *Pour μ-presque tout x dans X,*

$$\exists \lim_{n\to\infty} \frac{\log_2(R(x,n))}{n} = h(X, \mathcal{A}, \mu, T, \mathcal{P}).$$

7 Conclusion

Outre l'éclairage qu'ils apportent à l'étude des billards, des échanges d'intervalles, des systèmes symboliques et des transformations mesurables de $\mathbb{T}$, les deux théorèmes que nous venons de présenter répondent à une question naturelle concernant le temps de retour de Poincaré.

Pour certains exemples de systèmes d'entropie nulle, le temps de retour de Poincaré est lié à l'approximation diophantienne.

Michael D. Boshernitzan a introduit la notion suivante : un sous-shift minimal (X, S) est dit *prédictiblement récurrent* si pour deux points d'un même cylindre, les temps de retour sont proportionnels i.e. s'il existe une constante $C \geq 0$ telle que

$$(\forall (x,y) \in X^2)(\forall n \in \mathbb{N}^*)(P_n(x) = P_n(y) \Rightarrow R_n(x) \leq C R_n(y))$$

(voir paragraphe 6.3 pour les définitions).
De tels systèmes généralisent les sous-shifts linéairement récurrents et sont uniquement ergodiques. Ils semblent à priori intéressants pour étudier l'induction des échanges d'intervalles. Si $T = R_\varphi$ est la rotation d'angle φ irrationnel (voir paragraphe 2.1), on pose $\mathcal{P} = \chi_{[0,\{\varphi\}[}$ et dans ce cas, $(\overline{w(\mathbb{T})}, S)$ est linéairement récurrent si et seulement si il est prédictiblement récurrent si et seulement si les quotients partiels de φ sont bornés.

Références

[Ba] A. Baker, *Transcendental number theory*, Cambridge university press, (1975).

[Bo1] M. Boshernitzan, *A condition for unique ergodicity of minimal symbolic flows*, Ergodic theory and dynamical systems, 12 (1992), p 425-428.

[Bo2] M. Boshernitzan, *Quantitative recurrence results*, Inventiones mathematicae, 113 (1993), p 617-631.

[BS] L. Barreira, B. Saussol, *Hausdorff dimension of measures via Poincaré recurrence*, Communication in mathematical physics, 219 (2001), p 443-463.

[Ca] J.W.S. Cassels, *An introduction to diophantine approximation*, Cambridge university press, (1957).

[CFS] I.P. Cornfeld, S.V. Fomin, Ya.G. Sinai, *Ergodic theory*, Springer, (1982).

[CT] T. Cover, J. Thomas, *Elements of information theory*, Wiley, (1991).

[Di] P.G.L. Dirichlet, *Verallgemeinerung eines Satzes aus der Lehre von den Kettenbrüchen nebst einigen Anwendungen auf die Theorie der Zahlen*, (1842).

[Du] F. Durand, *Linearly recurrent subshifts*, Research report 98-02, IML (1998).

[Fa1] K.J. Falconer, *The geometry of fractal sets*, Cambridge tracts in mathematics, (1985).

[Fa2] K.J. Falconer, *Fractal geometry, mathematical fundations and applications*, Wiley, (1990).

[FGM] S. Ferenczi, L. Gaubert, T. Monteil *Sur l'unique ergodicité des échanges d'intervalles : le point de vue dynamique symbolique*, Notes de cours de DEA, en cours de rédaction.

[Gh] E. Ghys, *Variations autour du théorème de récurrence de Poincaré*, Le journal de maths de élèves, 1 (1994), p 3-12.
consultable sur : `http ://www.ens-lyon.fr/JME/Vol1Num1/artEGhys/artEGhys.html`

[Kr] U. Krengel, *Ergodic theorems*, de Gruyter Studies in mathematics, 6 (1985).

[OW] D.S. Ornstein, B. Weiss, *Entropy and data compression shemes*, IEEE transactions on information theory, 39 (1993), p 78-83.

[Po] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, tome 3, Gauthier-Villars, (1899).

[Ru] D. Ruelle, *Hasard et chaos*, Odile Jacob, (1991), p 150-154.

[Sc] W.M. Schmidt, *Diophantine approximation*, Lecture notes in mathematics, 785 (1980).

[Ta] S. Tabachnikov, *Billiards*, Panoramas et Synthèses, 1 (1995).

Annexe : Mesure et dimension de Hausdorff

Cette annexe s'appuie sur [Fa1]. Elle sert surtout à fixer les notations.

Soit (X, d) un espace métrique et α un réel positif. Soit A une partie de X.
Pour $\delta > 0$, on pose :

$$H_{\alpha,\delta}(A) \stackrel{\text{def}}{=} \inf\{\sum_{i\in\mathbb{N}} \operatorname{diam}(A_i)^\alpha \mid (A_i)_{i\in\mathbb{N}} \text{ est un } \delta\text{-recouvrement de } A\}.$$

Puis on définit une mesure extérieure sur X par

$$H_\alpha(A) \stackrel{\text{def}}{=} \lim_{\delta\to 0} H_{\alpha,\delta}(A) = \sup_{\delta>0} H_{\alpha,\delta}(A).$$

On appelle H_α la α-mesure de Hausdorff relative à d sur X.
La σ-algèbre $\mathcal{A}$ des ensembles mesurables pour H_α contient les boréliens de (X, d).
$H_\alpha(A)$ ne change pas si on remplace "δ-recouvrement" par "δ-recouvrement ouvert" ou "δ-recouvrement fermé" dans la définition de $H_{\alpha,\delta}(A)$.
Si $X = \mathbb{R}^n$ $(n \geq 1)$, $H_\alpha = \mu/c_n$ où μ désigne la mesure de Lebesgue et c_n le volume de la boule euclidienne de diamètre 1 ($c_1 = 1$ donc $H_1 = \mu$ sur $\mathbb{R}$).

Si $A \subset X$, l'application $\begin{pmatrix} \mathbb{R}_+ & \longrightarrow & [0,\infty] \\ \alpha & \longmapsto & H_\alpha(A) \end{pmatrix}$ est décroissante.
Si α est tel que $H_\alpha(A) \in \mathbb{R}_+^*$, alors $(\forall\beta < \alpha)(H_\beta(A) = \infty)$ et $(\forall\beta > \alpha)(H_\beta(A) = 0)$.
Ainsi, on peut définir la dimension de Hausdorff de A par :

$$\dim(A) \stackrel{\text{def}}{=} \sup\{\alpha \in [0,\infty] \mid H_\alpha(A) = \infty\} = \inf\{\alpha \in [0,\infty] \mid H_\alpha(A) = 0\} \in [0,\infty].$$

Si M est une sous-variété de dimension p de $\mathbb{R}^n$ $(n \geq 1)$, alors $\dim(M) = p$.

Enfin, si μ est une mesure borélienne sur X telle que $\mu(X) = 1$, on pose

$$\dim(\mu) \stackrel{\text{def}}{=} \min\{\dim(A) \mid A \in \mathcal{B}(X) \text{ et } \mu(A) = 1\}.$$

Pour des calculs explicites de la dimension de Hausdorff, on pourra consulter [Fa2].