Information, entropie, mesures feuilletées

François Ledrappier, LPSM, CNRS, Paris

Séminaire Betty B.

26 nov. 2021

Soit $\alpha \in \mathbb{R}/\mathbb{Z}$ irrationnel. Alors $\{2^n 3^m \alpha (\text{mod}.1), n, m \in \mathbb{N}\}$ est dense dans $\mathbb{R}/\mathbb{Z}$ [H. Furstenberg 67].

Question (Furstenberg 67): *Cette suite est-elle équirépartie?*

Cette question, toujours ouverte, s'est révélée extrêmement féconde. La réponse partielle obtenue se reformule en termes d'entropie des systèmes dynamiques et a éclairé de nombreux problèmes analogues.

1. **L'information de Shannon**
2. **L'entropie en théorie ergodique (Kolmogoroff)**
3. **Solénoïdes; problème de Furstenberg**
4. **Automorphismes de tores**
5. **Compléments**

Comme souvent, les conversations avec Jean-Paul Thouvenot m'ont beaucoup aidé à préparer cette présentation.

1. L'information de Shannon (1948)

$(\Omega, \mathcal{A}, m)$ un espace de probabilité, $P = (P_1, P_2, \ldots)$ une partition dénombrable de l'espace Ω en ensembles $\mathcal{A}$-mesurables.

Information (ou entropie) de P:

$$H(P) := -\sum_i m(P_i) \ln m(P_i) \quad (\text{avec } \ 0 \log 0 = 0.)$$

Propriétés: $0 \leq H(P) \leq \ln \#P$, avec

$H(P) = 0 \quad \Longleftrightarrow \quad m(P_i) = 0$ ou 1 pour tout i.

$\#P < +\infty$ et $H(P) = \ln \#P \iff m(P_i) = \frac{1}{\#P}$ pour tout i.

Interprétation: Les P_i sont les réponses possibles à une question et $H(P)$ mesure la quantité d'information apportée par la réponse.

Si les partitions $P = (P_1, \ldots)$ et $Q = (Q_1, \ldots)$ sont indépendantes (i.e. $m(P_i \cap Q_j) = m(P_i)m(Q_j)$), alors, en notant $P \vee Q$ la partition en $P_i \cap Q_j$,

$$H(P \vee Q) \ = \ H(P) + H(Q).$$

Plus généralement:

$$\begin{aligned}
H(P \vee Q) - H(P) \quad &= \quad \sum_{i,j} m(P_i \cap Q_j) \ln \frac{m(P_i)}{m(P_i \cap Q_j)} \\
&= \quad \sum_i m(P_i) \sum_j \frac{m(P_i \cap Q_j)}{m(P_i)} \ln \frac{m(P_i)}{m(P_i \cap Q_j)} \\
&= \quad \sum_i m(P_i) H_{m(.|P_i)}(Q) \ =: \ H(Q \big| P)
\end{aligned}$$

$H(Q \big| P)$ représente l'information apportée par la question Q si on connaît déjà la réponse à la question P.

Entropie conditionnelle: remplacer P par une sous-σ-algèbre $\mathcal{B}$.

$\mathcal{A}$ dénombrablement engendrée: alors, il existe des *mesures conditionnelles* $m_x^{\mathcal{B}}$

Pour tout $A \in \mathcal{A}$, $X \mapsto m_x^{\mathcal{B}}(A)$ est $\mathcal{B}$- mesurable.

Pour tous $A \in \mathcal{A}, B \in \mathcal{B}$, $m(A \cap B) = \int_B m_x^{\mathcal{B}}(A)\, dm(x)$.

Exemple: $\mathcal{B} = \mathcal{B}_P$ faite d'unions d'éléments d'une partition P

Pour m-presque tout x, $m_x^{\mathcal{B}_P} = m(.|P_i)$ si $x \in P_i$.

Entropie conditionnelle: $H(Q\big|\mathcal{B}) := \int H_{m_x^{\mathcal{B}}}(Q)\, dm(x)$.

$H(Q\big|\mathcal{B})$ représente l'information apportée par la question Q si on connaît déjà tout $\mathcal{B}$.

Propriétés: $H(Q\big|\mathcal{B}) := \int H_{m_x^{\mathcal{B}}}(Q)\,dm(x).$

Si $\mathcal{B}$ est dénombrablement séparée, alors pour m-presque tout x, la mesure $m_x^{\mathcal{B}}$ est portée par l'atome de $\mathcal{B}$ contenant x.

$0 \le H(Q\big|\mathcal{B}) \le H(Q).$

$H(Q\big|\mathcal{B}) = 0 \iff \mathcal{B}_Q \subset \mathcal{B}; \quad H(Q\big|\mathcal{B}) = H(Q) \iff Q \perp \mathcal{B}.$

$H(P \vee Q\big|\mathcal{B}) \;=\; H(P\big|\mathcal{B}) + H(Q\big|\mathcal{B}_P \vee \mathcal{B}).$

En particulier, si $\mathcal{B}_P \vee \mathcal{B} = \mathcal{B}_Q \vee \mathcal{B}$, alors $H(P\big|\mathcal{B}) = H(Q\big|\mathcal{B})$.

$\mathcal{B}_n \searrow_{n\to\infty} \mathcal{B}_\infty \Rightarrow H(Q\big|\mathcal{B}_n) \nearrow H(Q\big|\mathcal{B}_\infty),$

$\mathcal{B}_n \nearrow_{n\to\infty} \mathcal{B}_\infty$ et $H(Q\big|\mathcal{B}_1) < +\infty \Rightarrow H(Q\big|\mathcal{B}_n) \searrow H(Q\big|\mathcal{B}_\infty)$

2. Théorie ergodique; entropie de Kolmogoroff (1958)

Donnée: $(X, \mathcal{A}, m; T)$, $(X, \mathcal{A}, m)$ espace de probabilité, $T : X \to X$ measurable et *préservant la mesure:* $\forall A \in \mathcal{A}, m(T^{-1}A) = m(A)$.

σ-algèbre des invariants $\mathcal{I} := \{I, I \in \mathcal{A}, m(I \Delta T^{-1} I) = 0\}$.

Définition $(X, \mathcal{A}, m; T)$ est dit *ergodique* si $\forall I \in \mathcal{I}, m(I) = 0$ ou 1.

Décomposition ergodique On peut écrire

$$m = \int m_x \, dm_{\mathcal{I}}(x)$$

où $m_{\mathcal{I}}$ est une mesure sur $(X, \mathcal{I})$ et, pour presque tout x, m_x est une mesure préservée par T et ergodique pour T.

Non-démonstration: Il semble suffire de prendre les mesures conditionnelles de m par rapport à $\mathcal{I}$. $m_x(T^{-1}A) = m_{Tx}(A)$ p.p. par unicité, $= m_x(A)$ car $x \mapsto m_x \in \mathcal{I}$. Mais comme la σ-algèbre $\mathcal{I}$ n'est PAS dénombrablement séparée, montrer que m_x est ergodique pour p.t. x n'est pas banal (la démonstration utilise le théorème ergodique ponctuel).

Exemples

Rotation: $(X, \mathcal{A}, m) = (\mathbb{R}/\mathbb{Z}, Borel, dt),\ R_\alpha(t) = t + \alpha \ (\text{mod}.1)$.
Ergodique ssi α est irrationnel.

Solénoïde:
$p \in \mathbb{N}_{>0},\ (X, \mathcal{A}, m) = (\mathbb{R}/\mathbb{Z}, Borel, dt),\ T_p(t) = p \times t \ (\text{mod}.1)$.
T_p préserve dt: $T_p^{-1}([a, b]) = \cup_{k=0}^{p-1}[\frac{a+k}{p}, \frac{b+k}{p}]$ et est ergodique.

Automorphismes de tores: $A \in SL_d(\mathbb{Z}), (X, \mathcal{A}, m) = (\mathbb{R}^d/\mathbb{Z}^d, Borel, Leb.),\ T_A(t) = At \ (\text{mod}.\mathbb{Z}^d)$. T_A préserve $Leb.$ $(\text{Det}A = 1)$, est ergodique ssi aucune valeur propre λ_i de A n'est racine de l'unité.

Décalage de Bernoulli: $(\Omega, \mathcal{C}, \mathbb{P})$ espace de probabilité,
$(X, \mathcal{A}, m) = \otimes^{\mathbb{N}}(\Omega, \mathcal{C}, \mathbb{P})_n$, ou bien $(X, \mathcal{A}, m) = \otimes^{\mathbb{Z}}(\Omega, \mathcal{C}, \mathbb{P})_n$.
T le décalage: si $\underline{\omega} = \{\omega_j, j \in \mathbb{N}\}$ (ou $\{\omega_j, j \in \mathbb{Z}\}$), $(T\underline{\omega})_j = \omega_{j+1}$.
T est ergodique ($\mathcal{I} \subset \cap_n \sigma(\omega_k, k \geq n)$ et loi 0-1 de Kolmogoroff).

Fait général: récurrence de Poincaré *Soit* $(X, \mathcal{A}, m; T)$ *comme plus haut,* $A \in \mathcal{A}, m(A) > 0$. *Alors pour m-presque tout* $x \in A$, *il existe une infinité d'entiers* n *tels que* $T^n x \in A$.

Démonstration: Soit $B := \cup_0^\infty T^{-n} A \setminus \cup_1^\infty T^{-n} A$. Par définition les $T^{-j} B, j \geq 0$, sont deux à deux disjoints et de même mesure. Nécessairement, $m(B) = 0$. $\square$

Corollaire financier: *Si une fonction mesurable* $f : X \to \mathbb{R}$ *satisfait* $f(x) \leq f(Tx)$ *pour m-presque tout* x, *alors* f *est constante m-presque partout.*

Entropie de Kolmogoroff $(X, \mathcal{A}, m; T)$, $P = \{P_1, P_2, \dots\}$ une partition de X en ensembles mesurables, $H(P) < +\infty$. On pose: pour $i < j \in \mathbb{N}$ (ou dans $\mathbb{Z}$) $P_i^j := \vee_{k=i}^{j-1} T^{-k} P$,
$h(P, T) := \lim_{n\to\infty} \frac{1}{n} H(P_0^n)$ (qui existe par sous-additivité)
$h(T) = h_m(T) = h(T, m) := \sup\{h(P, T), H(P) < +\infty\}$.
Remarque: Si T est inversible et bimesurable, $h(P, T^{-1}) = h(P, T)$

Fait important: (Sinaï) $|h(P, T) - h(Q, T)| \leq H(P \big| Q) + H(Q \big| P)$.
Démonstration: $H((P \vee Q)_0^n) - H(P_0^n) = H(Q_0^n \big| P_0^n) \leq n H(Q \big| P)$.

Cor. 1 $h(T) = \sup\{h(P, T), \#P < +\infty\}$.
Cor. 2 si $P_0^{+\infty} = \mathcal{A}$ (ou si $P_{-\infty}^{+\infty} = \mathcal{A}$), $h(T) = h(P, T)$.
Cor. 3 $h(T^n) = n\, h(T)$.
Cor. 4 $\mathcal{P} := \{A \in \mathcal{A} : \ h(\{A, X \setminus A\}, T) = 0\}$ est une σ-algèbre; elle s'appelle la σ-algèbre de Pinsker $\mathcal{P}$.

Exemples $h(R_\alpha) = 0$. Décalage de Bernoulli: $h(T) = H(\Omega, \mathbb{P})$.
On verra que $h_{Leb.}(T_p) = \ln p$, $h_{Leb.}(T_A) = \sum_1^n \max\{0, \ln \lambda_i\}$.

Formules quand $H(P), H(Q) < +\infty$

$h(P,T) = H(P\big|P_1^{+\infty})$ et, si T est inversible, $= H(P\big|P_{-\infty}^{0})$.

$h(P,T) = \lim_{n\to\infty} \frac{1}{n} H(P_0^n) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=o}^{n-1} H(T^{-k}P|P_{k+1}^n) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=o}^{n-1} H(P|P_1^{n-k})$.

$h(P \vee Q, T) = h(P,T) + h(Q\big|Q_1^{+\infty} \vee P_{-\infty}^{+\infty})$ (Abramov-Rokhlin).

$h(P \vee Q, T) = \lim_{n\to\infty} \frac{1}{n} H(P_0^n \vee Q_0^n) = \lim_{n\to\infty} \frac{1}{n} H(P_0^n) + \lim_{n\to\infty} \frac{1}{n} H(Q_0^n|P_0^n) =$

$\lim_{n\to\infty} \frac{1}{n} \sum_{k=o}^{n-1} \left(H(T^{-k}P|P_{k+1}^n) + H(T^{-k}Q|Q_{k+1}^n \vee P_0^n)\right)$

$= \lim_{n\to\infty} \frac{1}{n} \sum_{k=o}^{n-1} \left(H(P|P_1^{n-k}) + H(Q|Q_1^{n-k} \vee P_{-k}^{n-k})\right)$.

Si $P \subset \cap_n Q_n^{+\infty}$, alors $h(P,T) = 0$, i.e. $\mathcal{P} \cap Q_{-\infty}^{+\infty} \supset \cap_n Q_n^{+\infty}$.

$h(Q,T) = h(P \vee Q, T) = h(P,T) + h(Q|Q_1^{+\infty} \vee P_{-\infty}^{+\infty}) = h(P,T) + h(Q|Q_1^{+\infty}) = h(P,T) + h(Q,T)$.

Réciproquement, si $Q_0^{+\infty} = \mathcal{A}$, alors $\mathcal{P} = \cap_n Q_n^{+\infty}$.

$B \in \mathcal{P} \Longrightarrow H(Q|Q_1^{+\infty}) = H(Q\big|Q_1^{+\infty} \vee \{B, X \setminus B\}_{-\infty}^{+\infty})$ et donc $\{B, X \setminus B\}$ et Q sont relativement indépendants par rapport à $Q_1^{+\infty}$. De même, $\{B, X \setminus B\}$ et Q_0^k sont relativement indépendants par rapport à $Q_k^{+\infty}$, pour tout k, ce qui force $\{B, X \setminus B\} \subset Q_n^{+\infty}$ pour tout n. Vrai aussi si $Q_{-\infty}^{+\infty} = \mathcal{A}$. En particulier:

Théorème 1 $H(Q) < +\infty \Longrightarrow \cap_n Q_n^{+\infty} = \cap_k Q_{-\infty}^k = \mathcal{P}$.

3. Solénoïde Soit $p \in \mathbb{N}_{>1}$. Set

$T_p : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$, $T_p t := pt$ (mod.1).

$P := \{[k/p, (k+1)/p), k = 0 \dots p-1\}$,

$P_0^{+\infty} = Borel = \mathcal{A}$ ($P_0^{+\infty}$ sépare les points irrationnels de $\mathbb{R}/\mathbb{Z}$).

$\mathcal{B} := P_1^{+\infty}$ ne sépare pas les $\{t + k/p, k = 0 \dots p-1\}$.

Soit m une mesure de probabilité sur $\mathbb{R}/\mathbb{Z}$, g une fonction mesurable telle que les mesures cond. $m_t^{\mathcal{B}}$ sont, pour m-p.t. t,

$$(*) \quad m_t^{\mathcal{B}} = \sum_{k=0}^{p-1} g(t+k/p)\delta_{t+k/p}.$$

On a, m-p.p., $g > 0$ et $\sum_{k=0}^{p-1} g(t+k/p) = 1$. Si m est T_p-invariante,

$h_m(T_p) = \int H_{m_t^{\mathcal{B}}}(P)dm(t) = -\int \sum_{k=0}^{p-1} g(t+k/p) \ln g(t+k/p)dm(t)$.

En particulier, $h_m(T_p) \leq \ln p$.

Si m est la mesure de Lebesgue dt alors g est constante $1/p$ et $h_{dt}(T_p) = \ln p$. Réciproquement, on a

Théorème 2 *Si $h_m(T_p) = \ln p$, alors $m = dt$.*

Théorème 2 *Si $h_m(T_p) = \ln p$, alors $m = dt$.*

Démonstration: On a m-p.p. $-\sum_{k=0}^{p-1} g(t+k/p) \ln g(t+k/p) = \ln p$ et donc $g(t) = 1/p$. Donc, la mesure m est préservée par $R_{1/p}$.

Lemme *La seule mesure invariante par T_p et par $R_{1/p}$ est dt.*

Démonstration: On a $g(t) = 1/p$ m-p.p. et aussi, en notant $\mathcal{B}_n := P_n^{+\infty}$,

$$(*_n) \quad m_t^{\mathcal{B}_n} = \frac{1}{p^n} \sum_{k=0}^{p^n-1} \delta_{t+k/p^n},$$

si bien que, pour tout n, $(R_{1/p^n})_* m = m$. D'autre part, $\{\alpha : (R_\alpha)_* m = m\}$ est un sous-groupe fermé de $\mathbb{R}/\mathbb{Z}$. $\square$

Théorème 3 *Si $h_m(T_p) = 0$, alors $(R_{1/p})_* m$ et m sont singulières l'une par rapport à l'autre.*

Démonstration: On a m-p.p. $-\sum_{k=0}^{p-1} g(t+k/p) \ln g(t+k/p) = 0$ et donc, pour m-presque tout t, un seul des $g(t+k/p), k = 0, \ldots, p-1$ vaut 1, et les autres valent 0. Il existe un ensemble mesurable A t. q. $m(A) = 1$ et $m(A+1/p) = 0$. $\square$

Remarque On a aussi que, pour tous k, n, $(R_{k/p^n})_* m$ et m sont singulières l'une par rapport à l'autre. Mais $\{\alpha : (R_\alpha)_* m \perp m\}$ n'est pas fermé a priori.

Mesures invariantes

Théorème 2 est un *principe variationnel*: la mesure de Lebesgue est la mesure invariante d'entropie maximale; il permet de reconnaître la mesure de Lebesgue parmi les mesures T_p-invariantes par une propriété purement dynamique. Il y a beaucoup de mesures T_p-invariantes:

1. $(T_p)_*\delta_0 = \delta_0$.
2. pour $q \in \mathbb{N}, (p|q) = 1$, $A_q := \{k/q, k = 1, \ldots, q-1\}$ est fini et T_p-invariant, il y a d'autres mesures discrètes T_p-invariantes.
3. La mesure de Lebesgue, de fonction $g = 1/p$,
4. **Proposition** *Pour toute g continue, $g \geq 0$ et $\sum_{k=0}^{p-1} g(t + k/p) = 1$, il existe une mesure T_p-invariante μ avec les $\mu_t^{\mathcal{B}}$ données par* $(*)$.

Démonstration: Soit $\mathcal{L}_g$ l'opérateur de transfert sur $C(\mathbb{R}/\mathbb{Z})$:

$\mathcal{L}_g f(t) = \sum_{k=0}^{p-1} g(t/p + k/p) f(t/p + k/p)$.

On a $\mathcal{L}_g 1 = 1$ et donc $\mathcal{L}_g^*$ préserve les probabilités sur $\mathbb{R}/\mathbb{Z}$.

$\mathcal{L}_g(f \circ T_p) = f$: un point fixe μ de $\mathcal{L}_g^*$ est T_p-invariant.

$\mathcal{L}_g(f_1 f_2 \circ T_p) = f_2 \mathcal{L}_g(f_1)$: les mesures $\mu_t^{\mathcal{B}}$ sont données par $(*)$. $\square$

Problème de Furstenberg *Supposons que p et q ne sont pas les puissances d'un même entier. Alors la seule mesure diffuse invariante par à la fois T_p et T_q est Lebesgue.*

Remarque Soit μ une telle mesure, $h_\mu(T_p)/\ln p = h_\mu(T_q)/\ln q$. En effet, soit P une partition en intervalles de taille $1/pq$, m, n tels que $|\frac{m}{n} - \frac{\ln p}{\ln q}| \leq \frac{1}{n}$. La partition $\vee_{k=0}^{n-1} T_p^{-k} P$ (respectivement $\vee_{k=0}^{m-1} T_q^{-k} P$) est une partition en intervalles de taille $1/p^n q$ (resp. $1/pq^m$). Chaque intervalle d'une partition rencontre au plus $p^2 q^2$ intervalles de l'autre.

Réductions On peut supposer μ ergodique pour l'action de T_p, T_q (i.e. tout ensemble invariant par T_p ET par T_q est de μ-mesure 0 ou 1). On peut considérer $h_\mu(T_p) = 0$ ou bien $h_\mu(T_p) > 0$.

Théorème 4 [R. Lyons 88, D. Rudolph 90, B. Host 95, W. Parry 96] *On suppose $(p|q) = 1$. Alors, la seule mesure diffuse invariante par à la fois T_p et T_q et d'entropie positive est Lebesgue.*

Voir les notes de Matan Tal (arXiv:2010.05989) pour l'historique, des détails, des démonstrations complètes et des commentaires.

Démonstration, d'après D. Rudolph and W. Parry
Soit $\mu = \int \mu_t d\mu_{\mathcal{I}_p}(t)$ la décomposition de μ en mesures ergodiques pour T_p. Comme T_p et T_q commutent, $T_q\mu_t = \mu_{T_q t}$, T_q agit sur $\mathcal{I}_p$, préserve $\mu_{\mathcal{I}_p}$ et est ergodique. En particulier, la fonction $t \mapsto h_{\mu_t}(T_p)$ est cste p.p..
On considère la fonction g_p telle que telle que pour μ-p. tout t,
$\mu_t^{\mathcal{B}} = \sum_{k=0}^{p-1} g_p(t+k/p)\delta_{t+k/p}$.

Lemme *On suppose* $(p|q) = 1$. *Alors,* $g_p(T_q t) = g_p(t)\, \mu - p.p.$.

Démonstration Les opérateurs de transfert $\mathcal{L}_p$ et $\mathcal{L}_q$ associés à (g_p, T_p) et (g_q, T_q) commutent et donc $g_p(T_q t)g_q(t) = g_q(T_p t)g_p(t)\, \mu$-p.p.. Puisque $(p|q) = 1$, T_p est une bijection sur $T_q^{-1}y$, pour tout y. Il vient:
$\mathcal{L}_q(g_p^{-1}(t)) = \sum_{y \in T_q^{-1}t} g_q(y)g_p^{-1}(y) = \sum_{y \in T_q^{-1}t} g_q(T_p y)g_p^{-1}(T_q y) =$
$g_p^{-1}(t) \sum_{y \in T_q^{-1}t} g_q(T_p y) = g_p^{-1}(t)\ \mu$ -p.p. (par définition, $g_p > 0\, \mu$-p.p.).
Le lemme suit. $\square$

Soit $\mathcal{H}$ la σ-algèbre engendrée par g_p. Le lemme montre que $T_q^{-1}\mathcal{H} = \mathcal{H}$ et donc $\mathcal{H} \subset \mathcal{P}_q$ (la σ-algèbre de Pinsker de T_q).
D'après la Remarque,

$$\lim_{n\to\infty} \frac{1}{n} H(\vee_{k=0}^{n-1} T_p^{-k} P \big| \mathcal{H}) = \lim_{n\to\infty} \frac{1}{n} H(\vee_{k=0}^{n-1} T_p^{-k} P).$$

Il s'ensuit que $\mathcal{H} \subset \cap_n \vee_{k=n}^{\infty} T_p^{-k} P$ (même dém. que $\mathcal{P}_p \subset \cap_m P_m^{+\infty}$), en particulier que $\mathcal{H} \subset \vee_{k=1}^{\infty} T_p^{-k} P$ et donc pour presque tout $t \in \mathbb{R}/\mathbb{Z}$, ou bien $g_p(t) = 1$ ou bien $g_p(t + j/p') = 1/p'$, où $j = 0, 1, \ldots, p' - 1$ et p' divise p.
Comme g_p n'est pas $1\,\mu_t$-p.p. (Théorème 3), on en déduit qu'il existe p' tel que $\mu(E_{p'}) > 0$, où $E_{p'} = \{t : (R_{1/p'})_* \mu_t = \mu_t\}$. Par récurrence de Poincaré,
$\mu(\{t :$ il existe une infinité d'entiers n_k avec $T_p^{n_k} t \in E_{p'}\}) > 0$.
Pour ces points t, $(R_{(p')^{-k}})_* \mu_t = \mu_t\})\ > 0$ pour tout k et donc, $\mu_t = Leb.$. Par ergodicité de T_q sur $\mathcal{I}_p$, pour $\mu_{\mathcal{I}_p}$-presque tout t, μ_t est Lebesgue et donc μ aussi. $\square$

Host 95 montre un Théorème plus fort que le Théorème 4.

Théorème 5 *Soit μ invariante ergodique d'entropie positive pour T_p. Alors pour μ-presque tout t, la suite $\{q^k t\}_{k\geq 0}$ est équirépartie.*

Démonstration que Théorème 5 implique Théorème 4: Soit μ diffuse invariante par à la fois T_p et T_q et d'entropie positive. On peut supposer μ ergodique pour l'action de T_p, T_q. Pour presque tout x, la mesure μ_x de la décomposition ergodique de μ pour T_p est invariante ergodique d'entropie positive pour T_p. Pour presque tout x, d'après le Théorème 5, la suite $\{T_q^k t\}_{k\geq 0}$ est équirépartie pour μ_x-presque tout t. Donc pour μ-presque tout t, la suite $\{T_q^k t\}_{k\geq 0}$ est équirépartie. Comme μ est T_q-invariante, μ est Lebesgue. $\square$

La démonstration du Théorème 5 utilise des simplifications inattendues dans le calcul des coëfficients de Fourier des limites des $\frac{1}{n}\sum_{k=0}^{n-1}\delta_{T_q^k t}$, pour μ-presque tout t.

Certaines démonstrations s'étendent (A. Johnson 92 pour Rudolph 90 et M. Hochman & P. Schmerkin 11 pour Host 95) au cas plus général où p et q ne sont pas les puissances d'un même nombre.

4. Automorphismes de tores

Soit $A \in SL_d(\mathbb{Z})$ tel que $T_A : T_A(t) = At\ (\mathbb{Z}^d)$ est ergodique sur $(\mathbb{R}^d/\mathbb{Z}^d, Borel, Leb.)$. Soient $\lambda_i, i = 1, \ldots, k$ les logarithmes des valeurs absolues des valeurs propres de A, $E_i, i = 1, \ldots, k$ les sous-espaces propres correspondants, $m_i = \mathrm{Dim} E_i, i = 1, \ldots k$, les multiplicités. Si $\widetilde{t}, \widetilde{t'} \in \mathbb{R}^d$, et $\widetilde{t} - \widetilde{t'} \in E_i$, alors, quand $n \to +\infty$,

$$e^{n(\lambda_i^+ - o(1))}|\widetilde{t} - \widetilde{t'}| \leq |A^n\widetilde{t} - A^n\widetilde{t'}| \leq e^{n(\lambda_i^+ + o(1))}|\widetilde{t} - \widetilde{t'}|,$$

où $\lambda_i^+ := \max\{0, \lambda_i\}$. Si P est une partition finie de $\mathbb{T}^d$ en parallélépipèdes assez petits de côtés parallèles aux E_i, P_0^n est une partition en parallélépipèdes de côtés parallèles aux E_i et de taille comprise entre $e^{-n(\lambda_i^+ + o(1))}$ et $e^{-n(\lambda_i^+ - o(1))}$. ("Assez petit" pour que les atomes de $P \cap T_A^{-1} P$ soient connexes.) On en conclut

Proposition *Pour toute mesure μ T_A-invariante sur $\mathbb{T}^d$, $h_\mu(T_A) \leq \sum_0^k m_i \lambda_i^+$. De plus, $h_{Leb.}(T_A) = \sum_0^k m_i \lambda_i^+$.*

Le principe variationnel est encore vrai:

Théorème 6 [K. Berg 69] *La mesure de Lebesgue est la seule mesure d'entropie* $\sum_0^k m_i \lambda_i^+$.

Plus généralement,

Théorème 7 [R. Bowen 71] *Si* T *est un endomorphisme d'un groupe compact, la seule mesure d'entropie maximale est la mesure de Haar.*

Mesures feuilletées

Les feuilletages $\mathcal{W}_i$ définis par les projections sur $\mathbb{T}^d$ des plans parallèles aux espaces E_i dans $\mathbb{R}^d$ ne définissent pas un espace quotient dénombrablement séparé. Mais dans une carte locale, les feuilles locales définissent un espace quotient dénombrablement séparé d'un ouvert borné de $\mathbb{R}^d$. Soit μ une mesure sur $\mathbb{T}^d$. On peut définir des mesures conditionnelles sur les feuilles locales. Là ou les cartes se chevauchent, les mesures locales sur la même feuille sont proportionnelles.

En paramétrant la feuille $\mathcal{W}_i(t)$ par $\mathbb{R}^{m_i}$ avec l'origine en t, on obtient une famille mesurable définie presque partout $t \mapsto \mu_t^i$ de mesures sur $\mathbb{R}^{m_i}$ que l'on peut normaliser en imposant $\mu_t^i(B_{m_i}(0,1)) = 1$; dans chaque carte locale, les mesures conditionnelles de μ sur la feuille locale $t + \mathcal{O}_i$, où $\mathcal{O}_i$ est un ouvert borné de R^{m_i}, sont proportionnelles à μ_t^i. Si la mesure μ est T_A-invariante, alors il existe une fonction mesurable ρ_i telle que, pour μ-presque tout t, $(T_A)_*(\mu_t^i) \;=\; \rho_i(t)\,\mu^i_{T_A t}$.
Pour la mesure de Lebesgue, $\rho_i(t)$ est la constante $(\det(A\big|_{E_i}))^{-1}$.

Entropies partielles

Plus généralement, on peut considérer le feuilletage $\mathcal{W}_I$ associé à un espace $E_I := \oplus_{i\in I} E_i$ et, les mesures conditionnelles μ_t^I associées à une mesure μ et à ces feuilletages W_I.

Si la mesure μ est T_A-invariante et $\lambda_i > 0$ pour tout $i \in I$, on définit

$$h_I(T_A, \mu) \; = \; \lim_{n\to\infty} \frac{1}{n} \int H_{\mu_t^I}(P_0^n)\, d\mu(t).$$

La limite ne dépend pas de la normalisation des μ_t^I ni du choix de la partition P en parallélépipèdes assez petits. On a encore

Théorème 8 *On suppose que $\lambda_i > 0$ pour tout $i \in I$. Alors:*
$h_I(T_A, \mu) \leq \sum_{i\in I} \lambda_i$ avec égalité ssi μ est Lebesgue;
$h_I(T_A, \mu) = 0$ ssi pour μ-presque tout t, μ_t^I est la mesure de Dirac en t.
$h_\mu(T_A) = h_{I^+}(T_A, \mu)$, où $I^+ := \{i, \lambda_i > 0\}$.

Le résultat (sous la forme où les conditionnelles sur les feuilles locales sont absolument continues par rapport à Lebesgue) est général pour un difféo. de classe C^2 et une mesure μ (L.&L.-S.Young 85).

Partitions subordonnées à un feuilletage invariant dilatant.
Un point technique est la construction de partitions *subordonnées* à $\mathcal{W}_i$.

Une partition η de $(X, \mathcal{A}, m)$ est dite mesurable si l'espace quotient $(X/\eta, \mathcal{A}/\eta, m)$ est dénombrablement séparé. Une partition mesurable est dite subordonnée à $\mathcal{W}_I$ si
les éléments de η sont des ouverts bornés de $\mathcal{W}_I$ et
η est croissante: pour m-presque tout t, $T_A^{-1}\eta(T_A t) \subset \eta(t)$.

Proposition [L.&J.-M. Strelcyn, 82] *On suppose que $\lambda_i > 0$ pour tout $i \in I$. Alors $\mathcal{W}_I$ admet des partitions subordonnées.*

Dans le cas des automorphismes de tores, une fois que l'on a une partition subordonnée au feuilletage $\mathcal{W}_I$, la démonstration du théorème 8 est similaire à celles des théorèmes 2 (principe variationnel) et 3 (caractérisation de l'entropie nulle) pour les solénoïdes.

Un résultat de rigidité (A. Katok & R. Spatzier, 96/98)
Soient deux matrices $A, B, \in SL_3(\mathbb{Z})$ qui commutent, E_1, E_2, E_3 les directions propres et $\Lambda_1, \Lambda_2, \Lambda_3$ les formes linéaires sur $\mathbb{R}^2$ décrivant les log des valeurs propres dans les directions E_1, E_2, E_3. On suppose que les formes linéaires Λ_i ne sont pas rationnelles.

Théorème 9 *Soit μ une mesure de probabilité sur $\mathbb{R}^3/\mathbb{Z}^3$ invariante ergodique par A et par B. On suppose qu'il existe E_i, Λ_i et $(p,q) \in \mathbb{Z} \times \mathbb{Z}$ avec $\Lambda_i((1,0)) = \alpha > 0, \Lambda_i((0,1)) = -\beta < 0$ et $h_i(T_{A^pB^q}) \neq 0$. Alors μ est Lebesgue.*

Démonstration On forme le système $(\mathbb{R}/\mathbb{Z} \times \mathbb{R}^3/\mathbb{Z}^3, \overline{T}, \overline{\mu})$, où $\overline{\mu} := ds \times \mu$, et $\overline{T}(s,t) = (R_{\frac{\alpha}{\alpha+\beta}} s, T_A t)$ si $s \in (0, \frac{\beta}{\alpha+\beta})$, $(R_{\frac{\alpha}{\alpha+\beta}} s, T_B t)$ si $s \in (\frac{\beta}{\alpha+\beta}, 1)$.
On a $\overline{T}^n(s,t) = (R_{n\frac{\alpha}{\alpha+\beta}} s, T_{A^{k_1}B^{k_2}} t)$, avec $|\Lambda_i((k_1,k_2))|$ borné.
La mesure $\overline{\mu}$ est invariante par $\overline{T}$. Soit $\overline{\mu} = \int \overline{\mu}_x \, d\mu_{\mathcal{I}}(x)$ la décomposition ergodique de $\overline{\mu}$. L'action de $Id \times T_A$ et $Id \times T_B$ sur $(\mathcal{I}, \mu_{\mathcal{I}})$ est ergodique et les propriétés qualitatives des $\overline{\mu}_x$ sont p.p. constantes.

Soit $\overline{\mu}^i_{s,t}$ une famille de mesures conditionnelles sur $\mathbb{R}$ (normalisées par $\overline{\mu}^i_{s,t}(B^i(t,1)) = 1$) associées à E_i et $\overline{\mu}_{s,t}$, pour presque tout (s,t).

Lemme 1 *Pour presque tout (s,t), la mesure $\overline{\mu}^i_{s,t}$ est proportionnelle à la mesure de Haar dt' d'un sous-groupe fermé de $\mathbb{R}$.*

Démonstration Par le théorème de Lusin, on peut choisir, pour tout $\varepsilon > 0$, un compact $K \subset \mathbb{R}/\mathbb{Z} \times \mathbb{R}^3/\mathbb{Z}^3$ tel que $\overline{\mu}(K) \geq 1 - \varepsilon^2$ et que $(s,t) \mapsto \overline{\mu}^i_{s,t}$ est continue sur K. On considère (s,t) tel que $\overline{\mu}^i_{s,t}(K) \geq 1 - \varepsilon$, et

$X_{(s,t)} := ((s, t + E_i)) \cap K \cap \mathrm{support}(\overline{\mu}^i_{s,t})$.

Pour tout t' tel que $(s,t') \in X_{(s,t)}$, il existe une suite $n_j \to \infty$ telle que $s + n_j \frac{\alpha}{\alpha+\beta} - m_j \searrow s$, $\overline{T}^{n_j}(s,t) \in K$ et $\overline{T}^{n_j}(s,t) \to (s,t')$. Il s'ensuit que les $\overline{\mu}^i_{\overline{T}^{n_j}(s,t)}$ convergent vers $\overline{\mu}^i_{s,t'}$. Comme $\overline{T}^{n_j}(s,t) = (R_{n_j \frac{\alpha}{\alpha+\beta}} s, T_{A^{k_1} B^{k_2}} t)$, avec $k_1 = \left[\frac{n_j \alpha}{\alpha+\beta}\right]$, $k_2 = \left[\frac{n_j \beta}{\alpha+\beta}\right]$ et $\Lambda_i(k_1,k_2) \to 0$, on a $\overline{\mu}^i_{s,t'} = \overline{\mu}^i_{s,t}$.

D'autre part, $\overline{\mu}^i_{s,t'}$ s'obtient à partir de $\overline{\mu}^i_{s,t}$ par la translation de t à t' sur $\mathcal{W}^i$. Comme ε est arbitrairement petit, la mesure $\overline{\mu}^i_{s,t}$est invariante par translation sur son support dans $\mathbb{R}$, à renormalisation près. Les seules classes de mesures avec cette propriété sont les mesures de Haar sur les sous-groupes. $\square$

Lemme 2 *Ou bien pour presque tout* (s,t), $\overline{\mu}^i_{s,t}$ *est la mesure de Lebesgue sur* $\mathbb{R}$, *ou bien pour presque tout* (s,t), $\overline{\mu}^i_{s,t} = \delta_0$.

Démonstration Le lemme 1 dit que les mesures conditionnelles non normalisées $\mu^i_{s,t}$ sont ou bien de la forme $\tau^{t'}(s,t)dt'$, pour $\tau(s,t)$ réel, ou bien portées par $t + \mathbb{Z}v_i(t)$, pour $v_i(t) \in E_i$, ou bien δ_0. Par ergodicité, les mesures $\mu^i_{s,t}$ sont presque toutes de même type.

Si le type est Lebesgue, l'action de $T_{A^pB^q}$ de la feuille $\mathcal{W}_i(s,t)$ sur la feuille de $\mathcal{W}_i(\overline{S}(s,t))$ multiplie $\mu^i_{s,t}$ par $\tau(s,t)^{p\alpha+q\beta}$ avec $p\alpha + q\beta \neq 0$ (car $h_i(T_{A^pB^q}) \neq 0$); par récurrence de Poincaré, il faut que $\tau(s,t) = 1$ $\overline{\mu}$-p.p..

Si le type est discret porté par $\{t + \mathbb{Z}v_i(t)\}$, l'action de $T_{A^pB^q}$ multiplie $||v_i(t)||$ par $e^{p\alpha+q\beta} \neq 1$; ce n'est pas possible, également par récurrence de Poincaré.

Reste le cas où le type est δ_0. $\square$

Si pour presque tout (s,t), $\overline{\mu}^i_{s,t}$ est la mesure de Lebesgue sur $\mathbb{R}$, pour presque tout x la mesure $\overline{\mu}_x$ est invariante par les translations le long de la direction irrationnelle E_i. D'autre part, la mesure $\overline{\mu}_x$ se projette sur $\mathbb{R}/\mathbb{Z}$ en la mesure invariante par la rotation irrationnelle $R_{\frac{\alpha}{\alpha+\beta}}$, Donc pour presque tout (s,t), $\overline{\mu}_{s,t}$ est la mesure de Lebesgue sur $\mathbb{R}/\mathbb{Z}\times\mathbb{R}^3/\mathbb{Z}^3$. La mesure $\overline{\mu}$ est $ds\times Leb.$ et $\mu = Leb.$.
On suppose donc que pour presque tout (s,t), $\overline{\mu}^i_{s,t}$ est la mesure de Dirac en $\{0\}$. Pour tout $j,k\in\mathbb{Z}$, la σ-algèbre $\mathcal{I}$ est invariante par $Id.\times T_{A^jB^k}$ On peut définir l'entropie relative le long de $\mathcal{W}^i$ conditionnellement par rapport à $\mathcal{I}$ par:

$$h_i(Id\times T_{A^jB^k},\overline{\mu}_x) \;=\; \lim_{n\to\infty}\frac{1}{n}\int H_{\overline{\mu}^i_{s,t}}(P_0^n)\,d\overline{\mu}_x(s,t).$$

Puisque les conditionnelles $\overline{\mu}^i_{s,t}$ sont des mesures de Dirac, on a, par une version conditionnelle du Théorème 8, $h_i(T_{A^jB^k},\overline{\mu}_x)=0$ pour tout (j,k). Comme la dimension est 3, il existe (j_0,k_0) tel que $\Lambda_i(j_0,k_0)$ n'a pas de $\Lambda_{i'}(j_0,k_0), i'\neq i$, du même signe. Alors l'entropie de $Id\times T_{A^{j_0}B^{k_0}}$ conditionnelle par rapport à $\mathcal{I}$ s'annule et donc les autres conditionnelles $\overline{\mu}^{i'}_{s,t}, i'\neq i$ sont aussi des mesures de Dirac. Toutes les entropies des $Id.\times T_{A^jB^k}$ viennent de la σ-algèbre $\mathcal{I}$. Il nous reste à examiner la σ-algèbre $\mathcal{I}$.

Le facteur $\mathbb{R}/Z, R_{\frac{\alpha}{\alpha+\beta}}$ est ergodique. Si la σ-algèbre $\mathcal{I}$ est saturée par les cercles, elle est de la forme $Borel \times \mathcal{J}$, où $\mathcal{J}$ est une sous σ-algèbre de parties boreliennes de $\mathbb{R}^3/\mathbb{Z}^3$. Alors $\mathcal{J}$ est invariante par T_A et par T_B. Par ergodicité, la σ-algèbre $\mathcal{J}$ est la σ-algèbre grossière et donc la σ-algèbre $\mathcal{I}$ aussi.

Sinon, il existe un entier $p \neq 0$ et une application mesurable φ_p de $\mathbb{R}/\mathbb{Z}$ dans $\mathbb{R}^3/\mathbb{Z}^3$ tels que le graphe $(s, \varphi_p(s))$ soit invariant par $\overline{T}^p$. Alors, la réunion des $(Id. \times T_{A^j B^k})(s, \varphi_p(s))$ est un ensemble invariant par $(Id. \times T_{A^j B^k}), j, k \in \mathbb{Z}$ qui est une union dénombrable de graphes. L'entropie sur cet ensemble est nulle. En examinant tour à tour les différentes valeur de p, on obtient un ensemble invariant où l'entropie est nulle. Le reste de l'espace est invariant et saturé par des cercles et le raisonnement précédent s'applique: ou bien il est de $\overline{\mu}$-mesure 1 et alors μ est Lebesgue, ou bien il est de $\overline{\mu}$-mesure 0, mais alors l'entropie de tout $T_{A^j B^k}$ est nulle, contrairement à l'hypothèse.

Finalement, on a montré que si les les formes linéaires Λ_i ne sont pas rationnelles, et qu'il existe une forme Λ_i telle que $\Lambda_i((1,0)) = \alpha > 0, \Lambda_i((0,1)) = -\beta < 0$ et $h_i(T_{A^p B^q}) \neq 0$ pour un certain $(p,q) \in \mathbb{Z}^2$, alors la mesure μ est Lebesgue. $\square$

5. Compléments

Un livre de référence est sur la page personnelle de Manfred Einsiedler:

M. Einsiedler, E.Lindenstrauss & T. Ward, *Entropy in ergodic theory and homogeneous dynamics.*

Pour le problème de Furstenberg, et les démonstrations des théorèmes 4 et 5,

Matan Tal, *Furstenberg's Times 2, Times 3 Conjecture (a Short Survey)* (arXiv:2010.05989).

La construction des partitions subordonnées est dans

F. Ledrappier & J.-M. Strelcyn, *A proof of the estimation from below in Pesin's entropy formula.* Ergodic Theory Dynam. Systems, vol. 2, 203–219 (1982).

La forme complète du Théorème 9 est dans

A. Katok & R. J. Spatzier, *Invariant measures for higher-rank hyperbolic abelian actions.* Ergodic Theory Dynam. Systems, vol. 16 751–778 (1996).

Corrections. Ergodic Theory Dynam. Systems, vol. 18, 503–507 (1998).

Les mesures feuilletées sont apparues en dynamique dans le cas où la mesure invariante est Lebesgue dans Ya. G. Sinai, *Classical dynamic systems with countably-multiple Lebesgue spectrum. II.* Izv. Akad. Nauk SSSR Ser. Mat. vol. 30, 15–68 (1966). (La première construction de partitions subordonnées est aussi dans cet article.) G. A. Margulis en a construit de nouvelles dans *Certain applications of ergodic theory to the investigation of manifolds of negative curvature.* Functional Anal. Appl. 3 (1969), no. 4, 335–336. Elles sont associées à des *courants positifs dans le sens des feuilles* dans D. Ruelle & D. Sullivan *Currents, flows and diffeomorphisms,* Topology, vol.14, 319–327 (1975). Et présentées comme *fonctions transverses* chez A. Connes *Cyclic cohomology and the transverse fundamental class of a foliation.* Geometric methods in operator algebras (Kyoto, 1983), 52–144, Pitman Res. Notes Math. Ser., 123, Longman Sci. Tech., Harlow, 1986. La définition donnée ici est celle de F. Ledrappier & L.-S. Young *The metric entropy of diffeomorphisms I and II,* Ann. of Math. (2), vol.122, 509-574, (1985). Les entropies partielles sont introduites dans cet article.

Il y a d'autres actions algébriques de $\mathbb{Z}^d, d > 1$, avec des phénomènes de rigidité comparables. Les deux exemples précédents ont servi d'inspiration.

Le flot des chambres de Weyl

Soient $X = SL_3(\mathbb{R})/SL_3(\mathbb{Z})$ l'espace des réseaux de volume 1 dans R^3 et $\alpha(t, s)$ l'action de $\mathbb{R}^2$ donnée par l'action à gauche des matrices $\begin{pmatrix} e^{-r-s} & 0 & 0 \\ 0 & e^r & 0 \\ 0 & 0 & e^s \end{pmatrix}$. Cette action préserve le quotient m de la mesure de Haar sur X.

Théorème 10 *Soit μ une mesure invariante ergodique. On suppose que l'une des transformations $\alpha(t, s)$ est d'entropie strictement positive. Alors $\mu = m$.*

M. Einsiedler, A. Katok & E. Lindenstrauss. *Invariant measures and the set of exceptions to Littlewood's conjecture.* Ann. of Math. (2) vol. 164, 513–560 (2006).

La conjecture de Littlewood est que pour tous réels u, v, on a $\liminf_{n \to +\infty} n\{nu\}\{nv\} = 0$. L'observation est que pour des réels u, v, on a $\liminf_{n \to +\infty} n\{nu\}\{nv\} = 0$ si, et seulement si, l'ensemble $\alpha(t, s) \begin{pmatrix} 1 & 0 & 0 \\ u & 1 & 0 \\ v & 0 & 1 \end{pmatrix}$, pour $s, t \geq 0$ est non-borné dans X. La condition d'entropie positive se traduit en dimension:

Corollaire *L' ensemble des* $(u, v) \in \mathbb{R}^2$ *qui ne satisfont pas la conjecture de Littlewood a dimension de Hausdorff 0.*

Groupes totalement discontinus

Soit $X = (\mathbb{Z}/2\mathbb{Z}, +)^{\mathbb{Z}^2}$. C'est un groupe compact pour la loi d'addition coordonnée par coordonnée. Les décalages T vers la gauche et S vers le haut définissent une action de $\mathbb{Z}^2$ par automorphismes. Soit A une partie de $\mathbb{Z}^2$. L'ensemble X_A des points de X tels que, pour tous $(n, m) \in \mathbb{Z}^2$, $\sum_{(i,j) \in A_{n,m}} x_{i,j} = 0$ où $A_{n,m}$ est le translaté de A de n horizontalement et de m verticalement, est un sous-groupe fermé $\mathbb{Z}^2$-invariant. La mesure de Haar μ_A sur ce groupe est $\mathbb{Z}^2$-invariante. On dit qu'une mesure $\mathbb{Z}^2$-invariante est algébrique si elle est de cette forme. Un exemple populaire est donné par $L = \{(0, 0), (0, 1), (1, 0)\}$.

Proposition *Pour tout* $s, 0 \le s \le \ln 4$, *il existe une mesure* μ_s S, T*-invariante d'entropie* $h_{\mu_s}(T) = s$ *sur* X_L. *Ces mesures* μ_s *sont algébriques.*

(M. Einsiedler, *Invariant subsets and invariant measures for irreducible actions on zero-dimensional groups.* Bull. L. Math. Soc. vol. 36, 321–331 (2004).)

Question *Est-ce-que toute mesure (T, S)-invariante sur X_L d'entropie strictement positive est algébrique?*

J.-F. Quint (*Percolation on the three dot system*, Probab. Theory Relat. Fields, vol. 149, 49–60 (2010)) décrit les mesures $\mathbb{Z}^2$-invariantes, mais cette description en répond pas immédiatement à la question.

Couplages (joinings)

Il a été remarqué, par exemple par A. Katok, S. Katok & K. Schmidt (*Rigidity of measurable structure for $\mathbb{Z}^d$-actions by automorphisms*, Comment. Math. Helv. vol 77, 718–745 (2002)) que les techniques précédentes de rigidité permettaient aussi de décrire les *couplages* d'actions de $\mathbb{Z}^2$. Le champ, la portée et les applications de cette remarque ont été considérablement étendus par M. Einsiedler et E. Lindenstrauss; voir Menny Aka, *Joining classification and applications (after M. Einsiedler and E. Lindenstrauss)*, Séminaire Bourbaki 1186, novembre 2021.