Comptes rendus de l'Académie des sciences. Série 1, Mathématique

PROBABILITÉS. — *Entropie mesurée et partitions de l'unité.* Note de **Étienne Ghys, Rémi Langevin** et **Pawel Walczak,** présentée par René Thom.

Dans cette Note nous définissons l'entropie d'un opérateur linéaire $\mathscr{D}$ de $L^{\infty}(X, \mu)$. Lorsque $\mathscr{D}$ est associé à une transformation T d'un espace de probabilité $(X, \mathscr{A}, \mu)$ préservant la mesure μ, son entropie est égale à celle de T. Nous espérons que d'autres opérateurs, comme la diffusion le long des feuilles d'un feuilletage admettrait une entropie non triviale.

PROBABILITY THEORY. — Metric entropy and partitions of unity.

In this Note we define the entropy of a linear operator $\mathscr{D}$ *of* $L^{\infty}(X, \mu)$. *When* $\mathscr{D}$ *is associated to a transformation* T *of a probability space* $(X, \mathscr{A}, \mu)$ *preserving the measure* μ *its entropy is equal to the entropy of* T. *We hope that other operators, as diffusion along the leaves of a foliation will admit non trivial entropy.*

Kolmogorov définit l'entropie d'une transformation T d'un espace de probabilité $(X, \mathscr{A}, \mu)$ qui préserve la mesure μ. L'entropie mesure la façon dont T opère sur les partitions finies de X.

En remarquant que T opère aussi sur les partitions de l'unité sur X, nous sommes amenés à définir l'entropie d'un opérateur linéaire $\mathscr{D}$ de $L^{\infty}(X, \mu)$ tel que

$$(*)\quad \begin{cases} (1) & f \geqq 0 \Rightarrow \mathscr{D}(f) \geqq 0; \\ (2) & \mathscr{D}(\mathbf{1}) = \mathbf{1}; \\ (3) & \int f\,d\mu = \int \mathscr{D}(f)\,d\mu. \end{cases}$$

Le but de cette Note est de montrer que l'entropie de l'opérateur $\mathscr{D}$ associé à une transformation est précisément égale à l'entropie mesurée (quelquefois appelée entropie métrique) de T.

Remarquons que d'autres extensions de la notion d'entropie existent déjà : entropie associée à un graphe [4], à un « Dynamische Verbänd » [5].

A un feuilletage d'une variété riemannienne compacte correspond un semi-groupe d'opérateurs $\mathscr{D}_t$ associé au laplacien feuilleté (*voir* [2]). Lucy Garnett démontre qu'il existe des mesures de probabilité, appelées mesures harmoniques, laissées invariantes par $\mathscr{D}_t$. L'entropie de cet opérateur n'est pas infinie. Bien que nous ne sachions pas expliciter son calcul, nous espérons cependant que l'entropie mesurée ainsi associée à une mesure harmonique ne sera pas toujours nulle. Il serait alors intéressant, comme ce fut fait pour un homéomorphisme, de la comparer à l'entropie géométrique du feuilletage (*voir* [3]) analogue à l'entropie topologique d'un difféomorphisme.

Soit (X, μ) un espace mesuré, μ une mesure de probabilité.

Donnons une définition de l'entropie d'une fonction, généralisant l'entropie $p.\log p$ de la fonction indicatrice d'une partie A de mesure p de X, et telle que l'entropie d'une fonction constante soit nulle.

DÉFINITION 1. — Soit φ une fonction mesurable essentiellement bornée positive ou nulle; posons :

$$\varepsilon(\varphi) = -\int_X \varphi \log \int \varphi + \int_X \varphi \log \varphi.$$

$\varepsilon(\varphi)$ est positive ou nulle et bornée, on vérifiera

$$\varepsilon(\varphi + \psi) \leqq \varepsilon(\varphi) + \varepsilon(\psi).$$

0249-6291/86/03030251 \$ 2.00

DÉFINITION 1′ — Soit φ et ψ deux fonctions positives ou nulles de $L^\infty(X, \mu)$. L'entropie conditionnelle de φ connaissant ψ est définie par :

$$\varepsilon(\varphi \mid \psi) = -\int_X \varphi\psi \log \frac{\int \varphi\psi}{\int \psi} + \int_X \varphi\psi \log \varphi.$$

On vérifiera : $\varepsilon(\varphi \mid 1) = \varepsilon(\varphi)$.

DÉFINITION 2. — Soit $\Phi = \{\varphi_1, \ldots, \varphi_n\}$ une partition de l'unité mesurable sur X. L'entropie de Φ est définie par :

$$\varepsilon(\Phi) = \sum_{i=1}^{n} \varepsilon(\varphi_i).$$

DÉFINITION 2′ — Soient $\Phi = \{\varphi_1, \ldots, \varphi_n\}$ et $\Psi = \{\psi_1, \ldots, \psi_m\}$ deux partitions de l'unité mesurables sur X, l'entropie de Φ connaissant Ψ est définie par :

$$\varepsilon(\Phi \mid \Psi) = \sum_{\alpha=1}^{n} \sum_{i=1}^{m} \varepsilon(\varphi_\alpha \mid \psi_i).$$

Définissons encore le produit de deux partitions de l'unité

$$\Phi = \{\varphi_1, \ldots, \varphi_n\}, \qquad \Psi = \{\psi_1, \ldots, \psi_m\}$$

par

$$\Phi \vee \Psi = \{\varphi_\alpha . \psi_i ;\ \alpha = 1, \ldots, n;\ i = 1, \ldots, m\}.$$

DÉFINITION 3. — L'image d'une partition de l'unité $\Phi = \{\varphi_1, \ldots, \varphi_n\}$ par un opérateur $\mathscr{D}$ satisfaisant les trois conditions de l'introduction (*) est :

$$\mathscr{D}(\Phi) = \{\mathscr{D}(\varphi_\alpha);\ \alpha = 1, \ldots, n\}.$$

Si $A = \{A_1, \ldots, A_n\}$ est une partition mesurable de X on notera Φ_A la partition de l'unité

$$\Phi_A = \{\chi_{A_1}, \ldots, \chi_{A_i}\};$$

on a :

$$\varphi_{A \vee B} = \varphi_A \vee \varphi_B.$$

Démontrons maintenant quelques relations analogues à celles démontrées en théorie ergodique (*voir* par exemple [6] théorème 4.3).

LEMME 1. — *Soient* Φ *et* Ψ *deux partitions de l'unité sur* X; *soient* A *et* B *deux partitions de* X *et* $\mathscr{D}$ *un opérateur satisfaisant* (*). *On a :*

(*a*) $\varepsilon(\Phi \vee \Psi) = \varepsilon(\Phi) + \varepsilon(\Phi \mid \Psi)$;

(*b*) $\varepsilon(\Phi \mid \Psi) \leqq \varepsilon(\Phi)$;

(*c*) $\varepsilon(\Phi \vee \Psi) \leqq \varepsilon(\Phi) + \varepsilon(\Psi)$;

(*d*) $\varepsilon(\Phi \vee \Psi \mid \Phi_A) \leqq \varepsilon(\Phi \mid \Phi_A) + \varepsilon(\Psi \mid \Phi_A)$;

(*e*) $\varepsilon(\Phi \mid \Phi_A \vee \Phi_B) \leqq \varepsilon(\Phi \mid \Phi_A)$;

(*f*) $\varepsilon(\Phi \mid \Phi_A) \geqq 0$.

Démonstration. — (*a*) se démontre par un calcul direct;

(*b*) se déduit de la convexité de la fonction $f(t) = t \log t$ puisque, si

$$C_i = \int \psi_i \qquad \text{et} \qquad r_i = \left(\int \alpha\psi_i\right)\left(\int \psi_i\right)^{-1}$$

on a

$$f(\sum_i C_i Z_i) \leqq \sum_i C_i f_{(r_i)};$$

(*c*) se déduit de (*a*) et (*b*);

(*d*) soit $A = \{ A_k \}$ (*d*) s'obtient en appliquant l'inégalité (*c*) aux partitions $\Phi_k = \{ \varphi_\alpha | A_k \}$ et $\Psi_k = \{ \psi_i | A_k \}$, A_k étant muni de la mesure de probabilité conditionnelle

$$\mu_k = \frac{1}{\mu(A_k)} \mu | A_k;$$

(*e*) soit $A = \{ A_1, \ldots, A_m \}$, $B = \{ B_1, \ldots B_n \}$ et $\Phi = \{ \varphi_1, \ldots, \varphi_r \}$; $\alpha \in \{ 1, \ldots, r \}$ et $j \in \{ 1, \ldots, n \}$ étant fixés, posons :

$$c_{k,l} = \frac{1}{\mu(B_j)} \mu(B_j \cap B_k \cap A_l)$$

$$s_{k,l} = \frac{1}{\mu(B_k \cap A_l)} \int_{B_k \cap A_l} \varphi_\alpha.$$

La convexité de la fonction f implique l'inégalité :

$$f(\sum_{k,l} C_{k,l} \cdot s_{k,l}) \leqq \sum_{k,l} C_{k,l} f(s_{k,l})$$

ce qui démontre (*e*);

(*f*) se déduit de la positivité de $\varepsilon(\varphi | A, \mu | A)$.

Définition 4. — Pour toute partition de l'unité Φ sur X et tout opérateur $\mathscr{D}$ satisfaisant les conditions (∗), posons lorsque cela a un sens

$$\varepsilon(\mathscr{D}, \Phi) = \lim_{n \to \infty} \frac{1}{n} \varepsilon\left(\bigvee_{k=0}^{n-1} \mathscr{D}^k \Phi \right);$$

en particulier le point (*c*) implique que cette limite existe en utilisant le raisonnement du lemme 1.19 de [1] lorsque $\mathscr{D} = {}^{T}\mathscr{D}$, *cf.* infra.

Considérons maintenant une transformation T de $(X, \mathscr{A}, \mu)$ préservant la mesure de probabilité μ et notons ${}^{T}\mathscr{D}$ l'opérateur défini par ${}^{T}\mathscr{D}(\varphi) = \varphi \circ T$. Cet opérateur vérifie les conditions (∗).

Lemme 2. — *Soient Φ et Ψ deux partitions de l'unité sur* X; *on a :*

$$\varepsilon({}^{T}\mathscr{D}\Phi | {}^{T}\mathscr{D}\Psi) = \varepsilon(\Phi | \Psi).$$

Lemme 3. — *Soient Φ une partition de l'unité sur* X *et* A *une partition mesurable de* X. *L'opérateur ${}^{T}\mathscr{D}$ satisfait l'inégalité*

$$\varepsilon({}^{T}\mathscr{D}, \Phi) \leqq \varepsilon({}^{T}\mathscr{D}, \Phi_A) + \varepsilon(\Phi | \Phi_A).$$

Démonstration. — Le lemme 1 implique l'inégalité

$$\varepsilon\left(\bigvee_{k=0}^{n-1} {}^{T}\mathscr{D}^k \Phi \right) \leqq \varepsilon\left(\bigvee_{k=0}^{n-1} {}^{T}\mathscr{D}^k \Phi \right) \vee \left(\bigvee_{j=0}^{n-1} \mathscr{D}^j \Phi_A \right).$$

On a donc :

$$\varepsilon\left(\bigvee_{k=0}^{n-1} {}^{T}\mathscr{D}^k \Phi \right) \leqq \varepsilon\left(\bigvee_{k=0}^{n-1} {}^{T}\mathscr{D}^k \Phi_A \right) + \varepsilon\left(\bigvee_{k=0}^{n-1} {}^{T}\mathscr{D}^k \Phi \,\middle|\, \bigvee_{j=0}^{n-1} {}^{T}\mathscr{D}^j \Phi_A \right)$$

$$\leqq \varepsilon\left(\bigvee_{k=0}^{n-1} {}^{T}\mathscr{D}^k \Phi_A \right) + \sum_{k=0}^{n-1} \varepsilon\left({}^{T}\mathscr{D}^k \Phi \,\middle|\, \bigvee_{j=0}^{n-1} {}^{T}\mathscr{D}^j \Phi_A \right)$$

et puisque ${}^{\mathrm{T}}\mathcal{D}^j(\Phi_{\mathrm{A}})$ provient de la partition $\{\mathrm{T}^j\mathrm{A}_i\}$ on a

$$\varepsilon\left(\bigvee_{k=0}^{n-1} {}^{\mathrm{T}}\mathcal{D}^k\Phi\right) \leqq \varepsilon\left(\bigvee_{k=0}^{n-1} {}^{\mathrm{T}}\mathcal{D}^k\Phi_{\mathrm{A}}\right) + \sum_{k=0}^{n-1} \varepsilon({}^{\mathrm{T}}\mathcal{D}^k\Phi \,|\, {}^{\mathrm{T}}\mathcal{D}^k\Phi_{\mathrm{A}}) \leqq \varepsilon\left(\bigvee_{k=0}^{n-1} {}^{\mathrm{T}}\mathcal{D}^k\Phi_{\mathrm{A}}\right) + n\,\varepsilon(\Phi\,|\,\Phi_{\mathrm{A}})$$

ceci quel que soit n.

THÉORÈME. — *L'entropie mesurée de la transformation* T *est égale à l'entropie de l'opérateur associé* ${}^{\mathrm{T}}\mathcal{D}$.

Démonstration. — L'entropie mesurée $\varepsilon(\mathrm{T})$ de la transformation T est

$$\varepsilon(\mathrm{T}) = \sup \varepsilon({}^{\mathrm{T}}\mathcal{D}, \Phi_{\mathrm{A}})$$

où A décrit l'ensemble des parties mesurables de X (*voir* [6] p. 87).

L'inégalité $\varepsilon(\mathrm{T}) \leqq \varepsilon({}^{\mathrm{T}}\mathcal{D})$ en découle.

Soient maintenant Φ une partition de l'unité sur X et ε un nombre positif. Il existe une partition A de X telle que l'entropie conditionnelle $\varepsilon(\Phi\,|\,\Phi_{\mathrm{A}})$ soit inférieure à ε. Il suffit de prendre

$$\mathrm{A} = \left\{ \bigcap \varphi_\alpha^{-1}\left\{\frac{i\alpha}{n}, \frac{i\alpha+1}{n}\right\},\ 0 \leqq i_\alpha \leqq r-1 \right\}$$

pour n assez grand.

Le lemme 3 implique alors que

$$\varepsilon({}^{\mathrm{T}}\mathcal{D}, \Phi) \leqq \varepsilon({}^{\mathrm{T}}\mathcal{D}, \Phi_{\mathrm{A}}) \pm \varepsilon \leqq \varepsilon(\mathrm{T}) + \varepsilon$$

et donc que pour tout $\varepsilon > 0$ on a $\varepsilon({}^{\mathrm{T}}\mathcal{D}) \leqq \varepsilon(\mathrm{T}) + \varepsilon$. □

Reçue le 14 avril 1986.

RÉFÉRENCES BIBLIOGRAPHIQUES

[1] R. BOWEN, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Springer Verlag, Lecture Notes in Math., n° 470, 1975.

[2] L. GARNETT, Foliations, the ergodic theorem and brownian motion, *J. Funct. Anal.*, 51, 1983, p. 285-311.

[3] E. GHYS, R. LANGEVIN et P. WALCZAK, *Geometric entropy of foliations*, preprint.

[4] M. GROMOV, *On the entropy of holomorphic maps*, preprint.

[5] G. PALM, Entropie und eizenger in dynamischen Verbänden, *Wahrsch. Verw. Geb.*, 36, 1976, p. 27-45.

[6] P. WALTERS, *An introduction to ergodic theory*, Springer Verlag, 1982.

Département de Mathématiques, Université de Lille-I,
59655 *Villeneuve-d'Ascq Cedex.*