COLLOQUIUM MATHEMATICUM

VOL. 84/85 2000 PART 1

A NOTE ON THE ENTROPY OF A DOUBLY STOCHASTIC OPERATOR

BY

BRUNON KAMIŃSKI (TORUŃ)
AND JOSÉ DE SAM LAZARO (ROUEN)

Abstract. We investigate the properties of the entropy and conditional entropy of measurable partitions of unity in the space of essentially bounded functions defined on a Lebesgue probability space.

Introduction. The entropy theory of dynamical systems has been enriched by introducing in [1] the concept of the entropy of a doubly stochastic ($\mathcal{DS}$) operator acting on the Banach space of essentially bounded functions defined on a Lebesgue probability space. It is shown there that this entropy is an extension of the Kolmogorov–Sinai (KS) entropy, i.e. the entropy defined in [1] of any Koopman operator associated with a measure-preserving transformation is equal to the entropy of this transformation. It seems that the theory of the former entropy will be more difficult than that of the KS-entropy.

In this note we investigate the properties of the entropy and conditional entropy of measurable partitions of unity which form the basis of the entropy theory of $\mathcal{DS}$-operators. Interesting definitions of these concepts with some properties are given in [1]. Among other things, we sharpen and generalize these results.

First, we observe that the entropy of a partition of unity is subinvariant with respect to any $\mathcal{DS}$-operator and that its invariance for all partitions forces the $\mathcal{DS}$-operator to be a Koopman operator. The subinvariance allows us to define the mean entropy of a partition of unity with respect to any $\mathcal{DS}$-operator and we show that it coincides with the definition given in [1] for Koopman operators.

Applying natural absolutely continuous measures associated with a measurable partition of unity, we give a definition of the conditional entropy which is simpler than that given in [1] and reduces to it under natural assumptions. Our definition allows us to obtain new properties of the condi-

2000 *Mathematics Subject Classification*: Primary 37D05.

tional entropy. Among our properties there are generalizations of properties (a), (d)–(f) of Lemma 1 of [1] to arbitrary partitions of unity.

Applying a result of A. Iwanik ([2]) we show that the entropy of a $\mathcal{DS}$-operator is an invariant with respect to a natural conjugacy relation for $\mathcal{DS}$-operators.

The first author would like to acknowledge his debt to A. Iwanik for fruitful discussions concerning $\mathcal{DS}$-operators.

Result. Let $(X,\mathcal{B},\mu)$ be a Lebesgue probability space. For any $f \in L^1(X,\mu)$ we put

$$Ef = \int\limits_X f\,d\mu.$$

Let $\mathcal{P}$ denote the set of all finite measurable partitions of unity in $L^\infty(X,\mu)$, i.e. $\Phi = \{\varphi_1,\ldots,\varphi_m\} \in \mathcal{P}$ iff $\varphi_i \in L^\infty(X,\mu)$, $\varphi_i \ge 0$, $1 \le i \le m$ and $\varphi_1 + \ldots + \varphi_m = 1$.

If $\Phi,\Psi \in \mathcal{P}, \Phi = \{\varphi_1,\ldots,\varphi_m\}, \Psi = \{\psi_1,\ldots,\psi_n\}$ then the *join* of Φ and Ψ is the following partition of unity:

$$\Phi \vee \Psi = \{\varphi_i\psi_j : 1 \le i \le m,\ 1 \le j \le n\}.$$

Let $\mathcal{P}^+$ denote the subset of $\mathcal{P}$ consisting of $\Phi = \{\varphi_1,\ldots,\varphi_m\}$ with all φ_i strictly positive.

We say that $\Phi = \{\varphi_1,\ldots,\varphi_m\}$ is *induced by a measurable partition* $P = \{P_1,\ldots,P_m\}$ of X if $\varphi_i = \chi_{P_i}$, $1 \le i \le m$.

For $\Phi \in \mathcal{P}$ we denote by $|\Phi|$ the number of elements of Φ.

First we recall the definition of the entropy of $\Phi \in \mathcal{P}$ (cf. [1]) expressing it by the well known function

$$\eta(t) = \begin{cases} -t\log t, & t \in (0,\infty), \\ 0, & t = 0. \end{cases}$$

We have

$$\eta(st) = s\eta(t) + t\eta(s), \qquad s,t \in [0,\infty). \tag{1}$$

The *entropy* of $\Phi \in \mathcal{P}$ is the number

$$\varepsilon(\Phi) = \sum_{\varphi\in\Phi} \varepsilon(\varphi) \quad \text{where} \quad \varepsilon(\varphi) = \varepsilon_\mu(\varphi) = \eta(E\varphi) - E(\eta\circ\varphi).$$

Now we introduce the definition of the conditional entropy of one partition of unity with respect to another. Our definition is simpler than that in [1] and reduces to it under natural assumptions.

Let $\psi \in L^\infty(X,\mu)$ be such that $0 \le \psi \le 1$ and $\psi \ne 0$. Let μ_ψ be the measure absolutely continuous with respect to μ, defined by $d\mu_\psi = (E\psi)^{-1}\psi d\mu$. We denote by E_ψ the integration operator with respect to μ_ψ.

Let $\varepsilon_\psi(\varphi) = \varepsilon_{\mu_\psi}(\varphi)$, and let

$$\varepsilon(\varphi \mid \psi) = \begin{cases} E(\psi)\varepsilon_\psi(\varphi), & \psi \neq 0, \\ 0, & \psi = 0. \end{cases}$$

Now for $\Phi, \Psi \in \mathcal{P}$ we define, in the same way as in [1], the *conditional entropy* of Φ with respect to Ψ by

$$\varepsilon(\Phi \mid \Psi) = \sum_{\varphi \in \Phi} \sum_{\psi \in \Psi} \varepsilon(\varphi \mid \psi).$$

An easy calculation shows that

$$\varepsilon(\varphi \mid \psi) = -E(\varphi\psi) \log \frac{E(\varphi\psi)}{E(\psi)} + E(\varphi\psi \log \varphi), \qquad \varphi \in \mathcal{P}^+,\ \psi \in \mathcal{P}^+,$$

i.e. one obtains the definition of the conditional entropy in [1].

Properties of entropy and conditional entropy

PROPOSITION 1. *For any $\Phi \in \mathcal{P}$ we have*

$$0 \leq \varepsilon(\Phi) \leq \log |\Phi|.$$

The equality $\varepsilon(\Phi) = 0$ holds iff every $\varphi \in \Phi$ is a constant, and $\varepsilon(\Phi) = \log |\Phi|$ iff Φ is induced by a partition of X and $E\varphi = |\Phi|^{-1}, \varphi \in \Phi$.

P r o o f. The inequality $\varepsilon(\Phi) \geq 0$ has been shown in [1]. It is an immediate consequence of the Jensen inequality. Let now $\varepsilon(\Phi) = 0$ and $\varphi \in \Phi$, i.e. $\eta(E\varphi) = E(\eta \circ \varphi)$. For $a \in \mathbb{R}$ we define $A_a = \{\varphi < a\}$ and

$$c = \inf\{a \in \mathbb{R} : \mu(A_a) = 1\}.$$

Let $\varepsilon > 0$ be arbitrary. It is enough to show that

$$\mu(c - \varepsilon \leq \varphi < c + \varepsilon) = 1. \tag{2}$$

Obviously $\mu(A_{c+\varepsilon}) = 1$ and $\mu(\varphi \geq c - \varepsilon) > 0$. Now we check that the strict concavity of η implies that $\mu(\varphi \geq c - \varepsilon) = 1$, i.e. (2) is satisfied.

Define $A = A_{c-\varepsilon}$, $B = X \setminus A$ and $E_D = E_{\chi_D}, D \in \mathcal{B}$.

Suppose, on the contrary, that $\mu(A) > 0$. Since $\mu(B) > 0$ we have

$$\begin{aligned} \eta(E\varphi) &= \eta(\mu(A)E_A\varphi + \mu(B)E_B\varphi) \\ &> \mu(A)\eta(E_A\varphi) + \mu(B)\eta(E_B\varphi) = E(\eta \circ \varphi), \end{aligned}$$

which is a contradiction, i.e. (2) is satisfied.

The inequality $\varepsilon(\Phi) \leq \log |\Phi|$ easily follows from the Jensen inequality:

$$\varepsilon(\Phi) \leq |\Phi| \sum_{\varphi \in \Phi} \frac{1}{|\Phi|} \eta(E\varphi) \leq |\Phi| \eta\left(\frac{1}{|\Phi|}\right) = \log |\Phi|.$$

If $\varepsilon(\Phi) = \log|\Phi|$ then

$$E\Big(\sum_{\varphi\in\Phi} \eta\circ\varphi\Big) = \sum_{\varphi\in\Phi} \eta(E\varphi) - \log|\Phi| \le 0,$$

i.e. $E(\eta\circ\varphi) = 0$, $\varphi\in\Phi$. Hence, every $\varphi\in\Phi$ admits only the values 0 or 1, i.e. Φ is induced by a partition of X and so

$$\varepsilon(\Phi) = \sum_{\varphi\in\Phi} \eta(E\varphi).$$

Therefore, the equality $\varepsilon(\Phi) = \log|\Phi|$ and the strict concavity of η imply $E\varphi = |\Phi|^{-1}$, $\varphi\in\Phi$. ■

From Proposition 1 we get at once the following

COROLLARY. *For every $\Phi, \Psi \in \mathcal{P}$ we have $\varepsilon(\Phi\,|\,\Psi) \ge 0$, and $\varepsilon(\Phi\,|\,\Psi) = 0$ iff for any $\varphi\in\Phi, \psi\in\Psi$ the function φ is constant on the set $\{\psi > 0\}$.*

It follows quite easily that if $\Phi = \{\varphi_1,\ldots,\varphi_n\}$ then $\varepsilon(\Phi\,|\,\Phi) = 0$ iff there exists a measurable partition $\{B_1,\ldots,B_k\}$ of X and a partition $\{J_1,\ldots,J_k\}$ of $\{1,\ldots,n\}$ such that for any $1\le l\le k$ there exists $i\in J_l$ such that

$$\varphi_i = c_i\chi_{B_l} \quad \text{with} \quad \sum_{i\in J_l} c_i = 1.$$

DEFINITION 1. An operator $D : L^\infty(X,\mu) \to L^\infty(X,\mu)$ is said to be *doubly stochastic* ($\mathcal{DS}$) if it is linear, positive, $D1 = 1$ and $E\circ D = E$.

The well known examples of doubly stochastic operators are the Koopman operators and the conditional expectation operators.

For $\Phi = \{\varphi_1,\ldots,\varphi_m\} \in \mathcal{P}$ we put, as in [1],

$$D\Phi = \{D\varphi_1,\ldots,D\varphi_m\}.$$

We shall use the following inequality of Jensen type for $\mathcal{DS}$-operators.

JENSEN INEQUALITY. *If D is a $\mathcal{DS}$-operator in $L^\infty(X,\mu)$ and $\eta:\mathbb{R}\to\mathbb{R}$ is concave then for any $f\in L^\infty(X,\mu)$ we have*

$$\eta(Df) \ge D(\eta\circ f).$$

This can be proved in exactly the same way as the corresponding inequality for conditional expectations (cf. for example [3], II 47).

PROPOSITION 2. *For any $\Phi\in\mathcal{P}$ we have $\varepsilon(D\Phi) \le \varepsilon(\Phi)$. If the equality $\varepsilon(D\Phi) = \varepsilon(\Phi)$ holds for every $\Phi\in\mathcal{P}$ then D is a Koopman operator.*

P r o o f. The above inequality is an immediate consequence of the Jensen inequality for $\mathcal{DS}$-operators.

Now suppose that $\varepsilon(D\Phi) = \varepsilon(\Phi)$ for every $\Phi \in \mathcal{P}$. Since $\varepsilon(\varphi) \geq \varepsilon(D\varphi)$, $\varphi \in \Phi$, the above equality implies that

$$\varepsilon(\varphi) = \varepsilon(D\varphi), \qquad \varphi \in \Phi. \tag{3}$$

Substituting in (3) $\varphi = \chi_A$, $A \in \mathcal{B}$, one obtains $E(\eta(D\chi_A)) = 0$. Hence, $D\chi_A$ admits only two values 0 or 1, i.e. it is a characteristic function and so D is a Koopman operator. ∎

PROPOSITION 3. *For every $\Phi, \Psi \in \mathcal{P}$ we have*

$$\varepsilon(\Phi \mid \Psi) = \sum_{\varphi\in\Phi}\sum_{\psi\in\Psi} \eta(E(\varphi\psi)) - \sum_{\psi\in\Psi} \eta(E\psi) - \sum_{\varphi\in\Phi} E(\eta(\varphi)).$$

Proof. Applying (1), we get

$$\begin{aligned}
\varepsilon(\varphi \mid \psi) &= E(\psi)\varepsilon_\psi(\varphi) \\
&= E(\psi)\left[\eta\left(\frac{E(\varphi\psi)}{E(\psi)}\right) - \frac{E(\eta(\varphi)\psi)}{E(\psi)}\right] \\
&= E(\psi)\left[\frac{1}{E(\psi)}\eta(E(\varphi\psi)) + \eta\left(\frac{1}{E\psi}\right)E(\varphi\psi) - \frac{E(\eta(\varphi)\psi)}{E(\psi)}\right] \\
&= \eta(E(\varphi\psi)) - \frac{\eta(E(\psi))E(\varphi\psi)}{E(\psi)} - E(\eta(\varphi)\psi)
\end{aligned}$$

for any $\varphi \in \Phi$, $\psi \in \Psi$. Summing over all $\varphi \in \Phi$ and $\psi \in \Psi$ one obtains the desired result. ∎

A connection between the entropies $\varepsilon(D\Phi \mid D\Psi)$ and $\varepsilon(\Phi \mid \Psi)$, $\Phi, \Psi \in \mathcal{P}$, is very important in developing the entropy theory for $\mathcal{DS}$-operators.

It is well known that if D is a Koopman operator induced by a transformation of X which preserves μ then these entropies are equal.

Now we want to give examples which show that, in general, there is no connection between them.

EXAMPLE 1. Let D be the integration operator E and let $\Phi, \Psi \in \mathcal{P}$ be such that $\varepsilon(\Phi \mid \Psi) > 0$. Then $\varepsilon(D\Phi \mid D\Psi) = 0 < \varepsilon(\Phi \mid \Psi)$.

EXAMPLE 2. Let T be an ergodic transformation of a Lebesgue space and let $P = \{A, B\}$ be a measurable partition of X. Let $D = \frac{1}{2}(I + U_T)$ where I is the identity operator and let $\Phi = \{\chi_A, \chi_B\}$. Then $D\Phi = \{\psi_1, \psi_2\}$ where $\psi_1 = \frac{1}{2}(\chi_A + \chi_{T^{-1}A})$ and $\psi_2 = \frac{1}{2}(\chi_B + \chi_{T^{-1}B})$. Applying Proposition 3, one has $\varepsilon(\Phi \mid \Phi) = 0$ while $\varepsilon(D\Phi \mid D\Phi) > 0$ except in the trivial case when ψ_i is constant on the set $\psi_i > 0$, $i = 1, 2$.

PROPOSITION 4. *For every $\Phi, \Psi, \Lambda \in \mathcal{P}$ we have*

$$\varepsilon(\Phi \vee \Psi \mid \Lambda) = \varepsilon(\Phi \mid \Lambda) + \varepsilon(\Psi \mid \Lambda \vee \Psi).$$

P r o o f. It follows from Proposition 3 that

$$\varepsilon(\Phi \vee \Psi | \Lambda) = \sum_{\varphi,\psi,\lambda} \eta(E(\varphi\psi\lambda)) - \sum_{\lambda} \eta(E(\lambda)) - \sum_{\varphi,\psi} E(\eta(\varphi\psi)), \tag{4}$$

$$\varepsilon(\Phi \mid \Lambda) = \sum_{\varphi,\lambda} \eta(E(\varphi\lambda)) - \sum_{\lambda} \eta(E(\lambda)) - \sum_{\varphi} E(\eta(\varphi)), \tag{5}$$

$$\varepsilon(\Psi \mid \Lambda \vee \Phi) = \sum_{\varphi,\psi,\lambda} \eta(E(\varphi\psi\lambda)) - \sum_{\varphi,\lambda} \eta(E(\varphi\lambda)) - \sum_{\psi} E(\eta(\psi)), \tag{6}$$

where the summations are taken over all $\varphi \in \Phi$, $\psi \in \Psi$, and $\lambda \in \Lambda$.

Combining (4)–(6) and applying (1) one obtains the desired result. ■

Substituting, in Proposition 4, Λ consisting of the function $\lambda \equiv 1$ one gets at once the following

COROLLARY. *For every* $\Phi, \Psi \in \mathcal{P}$,

$$\varepsilon(\Phi \vee \Psi) = \varepsilon(\Phi) + \varepsilon(\Psi \mid \Phi).$$

As we have seen, in general $\varepsilon(\Phi \mid \Phi) > 0$, i.e. $\varepsilon(\Phi \vee \Phi) > \varepsilon(\Phi)$ by the above Corollary. The asymptotic behaviour of the sequence $\varepsilon_n = \varepsilon(\Phi_1 \vee \ldots \vee \Phi_n)$, $\Phi_n = \Phi$, $n \geq 1$, is not clear from the definition. However, we have $\varepsilon_n = o(n)$, $n \geq 1$ (see below).

The fact that $\varepsilon(\Phi \mid \Phi) > 0$ for some $\Phi \in \mathcal{P}$ implies that one cannot equip $\mathcal{P}$ with a metric analogous to the Rokhlin metric and so one has no useful approximation results as in the classical entropy theory (cf. [4]). An interesting approximation result for the conditional entropy is contained in the proof of the main theorem of [1].

PROPOSITION 5. *For every* $\Phi \in \mathcal{P}$ *and* $\delta > 0$ *there exists* $\Psi \in \mathcal{P}$ *induced by a partition of* X *such that* $\varepsilon(\Phi \mid \Psi) < \delta$.

An idea of the proof is given in [1]. It seemed to us that some details of it could be useful to the reader and therefore we give the proof.

P r o o f. Let $\Phi = \{\varphi_1, \ldots, \varphi_n\}$ and let $\lambda > 0$ be such that

$$|x - x'| < \lambda,\ x, x' \in [0, 1] \Rightarrow |\eta(x) - \eta(x')| < \frac{\delta}{2n}.$$

Let $Q = \{Q_0, \ldots, Q_{r-1}\}$ be a partition of $[0, 1)$ where $Q_i = [i/r, (i + 1)/r)$, $0 \leq i \leq r - 1$, $r > 1/\lambda$. We consider the partition P of X defined by

$$P = \varphi_1^{-1}(Q) \vee \ldots \vee \varphi_n^{-1}(Q)$$

and we denote by $\Psi \in \mathcal{P}$ the partition induced by P. We claim that for any $A \in P$,

$$\varepsilon_{\chi_A}(\varphi_i) = \eta(E_A\varphi_i) - E_A(\eta(\varphi_i)) < \delta/n, \quad 1 \leq i \leq n, \tag{7}$$

where E_A has the same meaning as in the proof of Proposition 1.

We have

$$A = \{i_k/r \le \varphi_k < (i_k+1)/r,\ 1 \le k \le n\}$$

where $0 \le i_k < r-1$, $1 \le k \le n$. Hence,

$$0 \le E_A\varphi_k - i_k/r < 1/r < \lambda$$

and so

$$|\eta(E_A\varphi_k) - \eta(i_k/r)| < \delta/(2n). \tag{8}$$

In the same way, for $x \in A$ one has

$$|\eta(\varphi_k(x)) - \eta(i_k/r)| < \delta/(2n)$$

and therefore

$$|E_A(\eta(\varphi_k)) - \eta(i_k/r)| < \delta/(2n). \tag{9}$$

Combining (9) with (10) one obtains

$$|\eta(E_A\varphi_k) - E_A\eta(\varphi_k)| < \delta/n.$$

By the Jensen inequality the last difference is non-negative, i.e. one gets (8). Thus,

$$\varepsilon(\Phi \,|\, \Psi) = \sum_{i=1}^{n} \sum_{A\in P} \mu(A)\varepsilon_{\chi_A}(\varphi_i) < \delta. \ \blacksquare$$

PROPOSITION 6. *For every $\Phi, \Psi, \Lambda \in \mathcal{P}$ we have*

$$\varepsilon(\Phi \,|\, \Psi \vee \Lambda) \le \varepsilon(\Phi \,|\, \Lambda).$$

Proof. We have

$$\begin{aligned}
\varepsilon(\Phi \,|\, \Psi \vee \Lambda) &= \sum_{\varphi,\psi,\lambda} E(\psi\lambda)\varepsilon_{\psi\lambda}(\varphi) \\
&= \sum_{\varphi,\psi,\lambda} E(\psi\lambda)(\eta(E_{\psi\lambda}(\varphi)) - E_{\psi\lambda}(\eta(\varphi))) \\
&= \sum_{\varphi,\lambda} E(\lambda) \sum_{\psi} \frac{E(\psi\lambda)}{E(\lambda)} \eta(E_{\psi\lambda}(\varphi)) - \sum_{\varphi} E(\eta(\varphi)).
\end{aligned}$$

Applying the Jensen inequality and (1) one obtains

$$\begin{aligned}
\sum_{\varphi,\lambda} E(\lambda) \sum_{\psi} \frac{E(\psi\lambda)}{E(\lambda)} \eta(E_{\psi\lambda}(\varphi)) &\le \sum_{\varphi,\lambda} E(\lambda)\eta\biggl(\sum_{\psi} \frac{E(\psi\lambda)}{E(\lambda)} E_{\psi\lambda}(\varphi)\biggr) \\
&= \sum_{\varphi,\lambda} E(\lambda)\eta\biggl(\frac{E(\varphi\lambda)}{E(\lambda)}\biggr)
\end{aligned}$$

$$= \sum_{\varphi,\lambda} E(\lambda)\left[\frac{1}{E(\lambda)}\eta(E(\varphi\lambda)) + \eta\left(\frac{1}{E(\lambda)}\right)E(\varphi\lambda)\right]$$
$$= \sum_{\varphi,\lambda} \eta(E(\varphi\lambda)) - \sum_{\lambda} \eta(E(\lambda)).$$

Applying now Proposition 3 we have

$$\varepsilon(\Phi \mid \Psi \vee \Lambda) \le \sum_{\varphi,\lambda} \eta(E(\varphi\lambda)) - \sum_{\lambda} \eta(E(\lambda)) - \sum_{\varphi} E(\eta(\varphi)) = \varepsilon(\Phi \mid \Lambda). \blacksquare$$

COROLLARY 1. *For every $\Phi, \Psi \in \mathcal{P}$ we have $\varepsilon(\Phi \mid \Psi) \le \varepsilon(\Phi)$, and equality holds iff Φ and Ψ are uncorrelated, i.e.*

$$E(\varphi\psi) = E(\varphi)E(\psi)$$

for all $\phi \in \Phi$ and $\psi \in \Psi$.

Proof. The above inequality is an obvious consequence of Proposition 6.

If Φ and Ψ are uncorrelated then we obtain at once the desired equality by Proposition 3 and (1).

Now suppose that $\varepsilon(\Phi \mid \Psi) = \varepsilon(\Phi)$. Using the inequality and the concavity of η, it follows from a straightforward computation that

$$\varepsilon(\varphi) = \sum_{\psi \in \Psi} E(\psi)\varepsilon_{\psi}(\varphi), \qquad \varphi \in \Phi,$$

i.e.

$$\eta(E(\varphi)) = \sum_{\psi \in \Psi} E(\psi)\eta\left(\frac{E(\varphi\psi)}{E(\psi)}\right)$$

and so

$$\eta\Big(\sum_{\psi \in \Psi} E(\psi)E(\varphi)\Big) = \sum_{\psi \in \psi} E(\psi)\eta\left(\frac{E(\varphi\psi)}{E(\psi)}\right).$$

Applying the strict concavity of η one obtains

$$E(\varphi\psi) = E(\varphi)E(\psi), \qquad \varphi \in \Phi,\ \psi \in \Psi,$$

i.e. Φ and Ψ are uncorrelated. $\blacksquare$

Hence and by the Corollary to Proposition 4, one gets

COROLLARY 2. *For every Φ and Ψ we have*

$$\varepsilon(\Phi \vee \Psi) \le \varepsilon(\Phi) + \varepsilon(\Psi),$$

and equality holds iff Φ and Ψ are uncorrelated.

Let now D be a doubly stochastic operator and let $\Phi \in \mathcal{P}$.

Definition 3. The quantity

$$\varepsilon(D,\Phi)=\limsup_{n\to\infty}\frac{1}{n}\varepsilon\Big(\bigvee_{k=0}^{n-1}D^k\Phi\Big)$$

is said to be the *mean entropy* of D with respect to Φ.

It would be interesting to know whether the above sequence is convergent.

If D is a Koopman operator then we obtain Definition 4 of [1].

Corollary 2 and Proposition 2 yield at once the following

Corollary 4. $\varepsilon(D,\Phi)\le\varepsilon(\Phi)$.

Definition 4. The quantity

$$\varepsilon(D)=\sup\{\varepsilon(D,\Phi):\Phi\in\mathcal{P}\}$$

is said to be the *entropy* of D.

The entropy of a $\mathcal{DS}$-operator is an invariant with respect to the following natural relation of conjugacy.

Let $(X,\mathcal{B},\mu)$ and $(Y,\mathcal{C},\nu)$ be Lebesgue probability spaces and let D and D' be doubly stochastic operators on the spaces $L^\infty(X,\mu)$ and $L^\infty(Y,\nu)$, respectively. Then D and D' are said to be *conjugate* if there exists a bijective linear transformation $V:L^\infty(Y,\nu)\longrightarrow L^\infty(X,\mu)$ such that V and V^{-1} are positive, $VD'=DV$ and

$$\int_X Vf\,d\mu=\int_Y f\,d\nu$$

for every $f\in L^\infty(Y,\nu)$.

It follows from Lemma 6 of [2] that V is induced by a point transformation $\varphi:X\to Y$, i.e. $(Vf)(x)=f(\varphi x)$. Of course φ is bijective a.e. and measure-preserving. Therefore, the fact that the entropy is an invariant easily follows from the definition.

It is shown in [1] that for every Koopman operator U_T induced by a measure-preserving transformation T one obtains

$$\varepsilon(U_T)=h(T)$$

where $h(T)$ is the Kolmogorov–Sinai entropy of T.

In particular, if T is the identity transformation, then $h(T)=0$ and therefore

$$\varepsilon(\Phi_1\vee\ldots\vee\Phi_n)=o(n),\qquad \Phi_k=\Phi,\ 1\le k\le n,\ n\ge1.$$

We further obtain immediately

Remark 1. *If D is a $\mathcal{DS}$-operator such that $D^k=D$ for some positive integer k then $\varepsilon(D)=0$.*

In particular the entropy of a conditional expectation operator equals 0.

Conjecture. Let G be a compact abelian group equipped with a normalized Haar measure λ and let μ be a Borel probability measure on G. The operator D_μ on $L^\infty(G,\lambda)$ defined by

$$(D_\mu f)(x) = \int_G f(x-y)\,\mu(dy)$$

is said to be the convolution operator determined by μ.

If μ is a Dirac measure concentrated at some $x_0 \in G$ then D_μ is, of course, the Koopman operator induced by the rotation $\tau x = x - x_0$ on G. It is well known that the entropy of τ equals 0, i.e. $\varepsilon(D_\mu) = 0$.

On the other hand, if μ is a Haar measure then $D_\mu = E$ and, therefore, $\varepsilon(D_\mu) = 0$.

It would be interesting to know whether $\varepsilon(D_\mu) = 0$ for any μ.

REFERENCES

[1] E. Ghys, R. Langevin et P. Walczak, *Entropie mesurée et partitions de l'unité*, C. R. Acad. Sci. Paris Sér. I 303 (1986), 251–254.

[2] A. Iwanik, *Pointwise induced semigroups of σ-endomorphisms*, Colloq. Math. 38 (1977), 27–35.

[3] P. A. Meyer, *Probabilités et Potentiel*, Hermann, Paris, 1966.

[4] V. A. Rokhlin, *Lectures on the entropy theory of transformations with invariant measure*, Uspekhi Mat. Nauk 22 (1967), no. 3, 3–56 (in Russian).

Faculty of Mathematics and Informatics
Nicholas Copernicus University
Chopina 12/18
87-100 Toruń, Poland
E-mail: bkam@mat.uni.torun.pl

Département de Mathématiques
Université de Rouen
Site Colbert
76130 Mont-Saint-Aignan, France
E-mail: lazaro@bayes.univ-rouen.fr

Received 16 July 1999;
revised 4 November 1999
(3792)