Beloslav Riečan

On the entropy and generators of dynamical systems

Applications of Mathematics, Vol. 41 (1996), No. 3, 161–168

Persistent URL: http://dml.cz/dmlcz/134319

ON THE ENTROPY AND GENERATORS OF DYNAMICAL SYSTEMS

BELOSLAV RIEČAN, Bratislava

(Received July 16, 1992)

Summary. Recently D. Dumitrescu ([4], [5]) introduced a new kind of entropy of dynamical systems using fuzzy partitions ([1], [6]) instead of usual partitions (see also [7], [11], [12]). In this article a representation theorem is proved expressing the entropy of the dynamical system by the entropy of a generating partition.

Keywords: entropy of dynamical systems, fuzzy partitions, entropy of generating partition

AMS classification: 28D20

0. INTRODUCTION

Given a probability space $(\Omega, \mathscr{S}, P)$ and a measure preserving transformation $T\colon \Omega \to \Omega$, Kolmogorov and Sinaj constructed an invariant $h(T)$ called the entropy of the dynamical system $(\Omega, \mathscr{S}, P, T)$. With help of the invariant they showed that there exist non-isomorphic Bernoulli schemes. Namely, if $(\Omega_1, \mathscr{S}_1, P_1, T_1)$ and $(\Omega_2, \mathscr{S}_2, P_2, T_2)$ are two isomorphic dynamical systems, then they have the same entropy $h(T_1) = h(T_2)$. Hence if in some case $h(T_1) \neq h(T_2)$, then $(\Omega_1, \mathscr{S}_1, P_1, T_1)$, $(\Omega_2, \mathscr{S}_2, P_2, T_2)$ cannot be isomorphic. (For references see [12].)

Of course, if $h(T_1) = h(T_2)$, then generally we cannot say anything about the isomorphism of the corresponding dynamical systems. Therefore we tried in [7] to construct a larger family of invariants $h_{\mathscr{G}}(T)$ by substituting set partitions by fuzzy set partitions, i.e., collections $\{f_1, \ldots, f_n\}$ of functions $f_i\colon \Omega \to \langle 0, 1\rangle$ such that $f_1 + \ldots + f_n = 1$.

The Dumitrescu approach ([4], [5] and also [12]) is a little more general. He does not assume the existence of a probability space and starts with a fuzzy probability space which is a set $\mathscr{F} \subset \langle 0,1\rangle^{\Omega}$ with a function $m\colon \mathscr{F} \to \langle 0,1\rangle$ satisfying some conditions. Of course, using a representation theorem from [3], we obtain

$$m(f) = \int f \,\mathrm{d}P$$

so that the Dumitrescu theory can be reduced to the case studied in [7]. Another approach using fuzzy sets theory has been developed by Markechová ([8], [9], see also Mesiar [10]).

Probably one of the most important results of the theory for practical purposes is the Kolmogorov-Sinaj theorem stating that

$$h(T) = h(T, \mathscr{A}),$$

whenever $\mathscr{A}$ is a partition generating the given σ-algebra. An analogue of this theorem is contained in [7]:

$$h_{\mathscr{G}}(T) \leqslant h(T, \mathscr{A}) + K_{\mathscr{G}},$$

where $K_{\mathscr{G}}$ is a constant depending on the family $\mathscr{G} \subset \langle 0,1\rangle^{\Omega}$ determining $h_{\mathscr{G}}(T)$. (Of course, if $\mathscr{G}$ consists of indicators χ_E only, then $K_{\mathscr{G}} = 0$.) Here $\mathscr{A}$ is a crisp partition generating the given algebra of sets. In this paper we present a variant of the Kolmogorov–Sinaj theorem in the form $h_{\mathscr{G}}(T) = h(T, \mathscr{A})$, of course, in the case that $\mathscr{A}$ is a fuzzy partition generating the given fuzzy σ-algebra $\mathscr{F}$.

1. Definitions

We assume that a set $\mathscr{F} \subset \langle 0,1\rangle^{X}$ is given satisfying the following conditions:

(i) $1_X \in \mathscr{F}$,

(ii) $f, g \in \mathscr{F} \Rightarrow g - \min(f,g) \in \mathscr{F}$,

(iii) $f, g \in \mathscr{F}, f + g \leqslant 1 \Rightarrow f + g \in \mathscr{F}$,

(iv) $f, g \in \mathscr{F} \Rightarrow f \cdot g \in \mathscr{F}$.

We say that a family satisfying the conditions stated above is a fuzzy algebra. Evidently $O_X \in \mathscr{F}$. It is not difficult to show that $\mathscr{F}$ is closed under the maximum and the minimum. Indeed, if $f, g \in \mathscr{F}$, then

$$f S_{\infty} g = \min(f + g, 1) = 1 - (1 - f - \min(1 - f, g)) \in \mathscr{F},$$
$$f T_{\infty} g = \max(f + g - 1, 0) = g - \min(1 - f, g) \in \mathscr{F},$$

and finally

$$\begin{aligned}\min(f,g) &= \max(f+\min(g-f,0),0)\\ &= \max(f+\min((1-f)+g,1)-1,0)\\ &= fT_\infty((1-f)S_\infty g) \in \mathscr{F},\\ \max(f,g) &= 1-\min(1-f,1-g) \in \mathscr{F}.\end{aligned}$$

Further we assume that a function $m\colon \mathscr{F} \to \langle 0,\infty)$ is given satisfying the following conditions:

(i) $f,g \in \mathscr{F}, f \leqslant g \Rightarrow m(f) \leqslant m(g)$,

(ii) $f,g,h \in \mathscr{F}, f = g \dot{+} h \Rightarrow m(f) = m(g) + m(h)$.

As an easy consequence of (i) and (ii) we obtain

(iii) $f \in \mathscr{F},\ f_i \in \mathscr{F}\ (i=1,\dots,n),\ f \leqslant \sum\limits_{i=1}^{n} f_i \Rightarrow m(f) \leqslant \sum\limits_{i=1}^{n} m(f_i)$.

Indeed, if $f,g,h \in \mathscr{F},\ f \leqslant g+h$, then also $f \leqslant \min(g+h,1) \in \mathscr{F}$, hence

$$\begin{aligned}m(f) \leqslant m(\min(g+h,1)) &= m(g) + m(\min(g+h,1)-g)\\ &= m(g) + m(\min(h,1-g)) \leqslant m(g) + m(h).\end{aligned}$$

The third part of our assumptions is concerned with a mapping $U\colon \mathscr{F} \to \mathscr{F}$ satisfying the following conditions:

(i) $f,g,h \in \mathscr{F},\ f = g+h \Rightarrow U(f) = U(g) + U(h)$,

(ii) $U(1_X) = 1_X$,

(iii) $f \in \mathscr{F} \Rightarrow m(U(f)) = m(f)$.

As a special case of the previous definition one can consider a mapping $U\colon \mathscr{F} \to \mathscr{F}$ induced by a transformation $T\colon X \to X$ by the formula $Uf(x) = f(T(x))$. In fact, this is the classical case.

As we have mentioned, a fuzzy partition is a finite collection $\mathscr{A} = \{f_1,\dots,f_n\}$ of members of $\mathscr{F}$ such that $\sum\limits_{i=1}^{n} f_i(x) = 1$ for all $x \in X$. If $\mathscr{A} = \{f_1,\dots,f_n\}$ is a fuzzy partition, then we define $U\mathscr{A} = \{Uf_1,\dots,Uf_n\}$. Evidently $U\mathscr{A}$ is a fuzzy partition, too. If $\mathscr{A} = \{f_1,\dots,f_n\}$, $\mathscr{B} = \{g_1,\dots,g_n\}$ are two partitions, then we put

$$\mathscr{A} \vee \mathscr{B} = \{f_i \cdot g_j\,;\ i=1,\dots,n,\ j=1,\dots,m\}.$$

Also $\mathscr{A} \vee \mathscr{B}$ is a partition. The entropy of the partition $\mathscr{A}$ is defined by the formula

$$H(\mathscr{A}) = \sum_{i=1}^{n} \varphi(m(f_i)),$$

where

$$\varphi(x) = -x \log x, \quad \text{if} \quad x > 0, \varphi(0) = 0.$$

The entropy of the dynamical system $(X, \mathscr{F}, m, U)$ is defined as follows. If $\mathscr{A}$ is a partition, then

$$h(U, \mathscr{A}) = \lim_{n \to \infty} \frac{1}{n} H(\mathscr{A} \vee U\mathscr{A} \vee \ldots \vee U^{(n-1)}\mathscr{A}).$$

(The existence of the limit follows by the subadditivity property ($a_{n+m} \leqslant a_n + a_m$) of the sequence $(a_n)_n$, where $a_n = H(\mathscr{A} \vee U\mathscr{A} \vee \ldots \vee U^{n-1}\mathscr{A})$.) Finally, if $\emptyset \neq \mathscr{G} \subset \mathscr{F}$, then

$$H_{\mathscr{G}}(U) = \sup \{h(U, \mathscr{A}); \mathscr{A} \subset \mathscr{G}, \mathscr{A} \text{ is a partition}\}.$$

2. Theorem

In order to be able to formulate the main result of the article we need the following notation. If $f, g\colon X \to \langle 0, 1\rangle$, then

$$f \bigtriangleup g = f(1-g) + g(1-f).$$

(Of course, if f, g are indicators, $f = \chi_A, g = \chi_B$, then $f \bigtriangleup g = \chi_{A \bigtriangleup B}$.)

Theorem. *If $\mathscr{C} \subset \mathscr{G}$ is a generator (i.e., such a partition that for every $\lambda > 0$ and every $f \in \mathscr{G}$ there is $g \in \bigcup_{i=0}^{\infty} U^i\mathscr{C}$ such that $m(f \bigtriangleup g) < \lambda$), then*

$$h_{\mathscr{G}}(U) = h(U, \mathscr{C}).$$

3. Proof

The main idea of the proof can be found in [2]. We shall divide it into a sequence of lemmas. Denote by $s(\mathscr{A})$ the fuzzy algebra generated by $\mathscr{A}$.

Lemma 1. *For a given partition $\mathscr{A} = \{f_1, \ldots, f_n\}$ and every $\delta > 0$ there exists $\lambda > 0$ such that for every family $\mathscr{B} = \{g_1, \ldots, g_n\}$ with $m(f_i \bigtriangleup g_i) < \lambda$ $(i = 1, 2, \ldots, n)$ there is a partition $\{h_1, \ldots, h_n\} \subset s(\mathscr{B})$ such that $m(f_i \bigtriangleup h_i) < \delta$ $(i = 1, 2, \ldots, n)$.*

P r o o f. Put

$$h_1 = g_1,$$

$$h_i = \min\left(g_i, 1 - \sum_{j=1}^{i-1} h_j\right), \ i = 2, \dots, n-1,$$

$$h_n = 1 - \sum_{j=1}^{n-1} h_j.$$

Evidently $\{h_1, \dots, h_n\}$ is a partition included in $s(\mathscr{B})$. Let $1 < i < n$. Then

$$(*) \qquad m(f_i \bigtriangleup h_i) \leqslant \sum_{j=1}^{i-1} m(h_j \bigtriangleup f_j) + m(f_i \bigtriangleup g_i).$$

Indeed, if $h_i(x) = g_i(x)$, then

$$h_i \bigtriangleup f_i(x) = f_i \bigtriangleup g_i(x).$$

If $h_i(x) = 1 - \sum_{j=1}^{i-1} h_j(x) \leqslant g_i(x)$, then

$$h_i(x)(1 - f_i(x)) \leqslant g_i(x)(1 - f_i(x)) \leqslant f_i \bigtriangleup g_i(x),$$

$$f_i(x)(1 - h_i(x)) = f_i(x) \sum_{j=1}^{i-1} h_j(x) \leqslant \sum_{j=1}^{i-1} (1 - f_j(x)) h_j(x) \leqslant \sum_{j=1}^{i-1} h_j \bigtriangleup f_j(x)$$

so that

$$h_i \bigtriangleup f_i \leqslant \sum_{j=1}^{i-1} h_j \bigtriangleup f_j + f_i \bigtriangleup g_i,$$

which implies $(*)$. Now let $i = n$. Then

$$(1 - h_n(x)) f_n(x) = \sum_{j=1}^{n-1} h_j(x) f_n(x) \leqslant \sum_{j=1}^{n-1} h_j(x)(1 - f_j(x)) \leqslant \sum_{j=1}^{n-1} h_j \bigtriangleup f_j(x),$$

$$(1 - f_n(x)) h_n(x) = \sum_{j=1}^{n-1} f_j(x) h_n(x) \leqslant \sum_{j=1}^{n-1} f_j(x)(1 - h_j(x)) \leqslant \sum_{j=1}^{n-1} f_j \bigtriangleup h_j(x),$$

hence

$$(**) \qquad m(f_n \bigtriangleup h_n) \leqslant \sum_{j=1}^{n-1} m(h_j \bigtriangleup f_j).$$

Since $m(f_i \bigtriangleup g_i) < \lambda$, we obtain by $(*)$ and $(**)$ that

$$m(h_i \bigtriangleup f_i) < 2^{i-1}\lambda \quad (i = 1, 2, \dots, n).$$

Therefore we can put $\lambda = \frac{\delta}{2^{n-1}}$. □

For any partitions $\mathscr{A} = \{f_1, \dots, f_n\}$, $\mathscr{B} = \{g_1, \dots, g_m\}$ we define the conditional entropy

$$H(\mathscr{A}|\mathscr{B}) = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ m(h_j)>0}}^{m} m(h_j)\varphi\left(\frac{m(f_i h_j)}{m(h_j)}\right).$$

Lemma 2. *For every $\varepsilon > 0$ there exists $\delta > 0$ such that $H(\mathscr{A}|\mathscr{B}) < \varepsilon$ for any partitions $\mathscr{A} = \{f_1, \dots, f_n\}$, $\mathscr{B} = \{h_1, \dots, h_n\}$ satisfying the condition $m(f_i \triangle h_i) < \delta$ $(i = 1, \dots, n)$.*

Proof. First choose $\delta_0 \in (0,1)$ such that $\varphi(t) < \frac{\varepsilon}{n}$ for every $t \notin (\delta_0, 1 - \delta_0)$ and put

$$\delta = \min\left\{\frac{\delta_0}{2} m(f_i);\ m(f_i) > 0\right\}.$$

Then

$$\begin{aligned}
m(f_i) \leqslant m(f_i \triangle h_i) + m(h_i) < \delta + m(h_i) < \delta_0 \frac{m(f_i)}{2} + m(h_i),\\
\frac{m(f_i)}{2} < m(f_i) - \delta_0 \frac{m(f_i)}{2} < m(h_i),\\
m(h_i) - m(f_i h_i) \leqslant m(h_i \triangle f_i) \leqslant \delta \leqslant \delta_0 m(h_i).
\end{aligned}$$

If we consider such an i that $m(h_i) > 0$, then

$$\frac{m(f_i h_i)}{m(h_i)} > 1 - \delta_0,$$

hence

$$\varphi\left(\frac{m(f_i h_i)}{m(h_i)}\right) < \frac{\varepsilon}{n},$$

$$H(\mathscr{A}|\mathscr{B}) = \sum_i \sum_j m(h_i)\ \varphi\left(\frac{m(f_i h_i)}{m(h_i)}\right) < \sum_i \sum_j m(h_i)\frac{\varepsilon}{n} = \sum_j \frac{\varepsilon}{n} = \varepsilon.$$

□

Lemma 3. *$h(U, \mathscr{A}) \leqslant h(U, \mathscr{C}) + H(\mathscr{A}|\mathscr{C})$ for any partitions $\mathscr{A}, \mathscr{C}$.*

Proof. Since $h(\mathscr{B} \vee \mathscr{D}) = H(\mathscr{B}) + H(\mathscr{D}|\mathscr{B})$ (see [5], [6]), we have

$$\begin{aligned}
H\Big(\bigvee_{i=0}^{n-1} U^i \mathscr{A}\Big) &\leqslant H\Big(\bigvee_{i=0}^{n-1} U^i \mathscr{A} \vee \bigvee_{j=0}^{n-1} U^j \mathscr{C}\Big)\\
&= H\Big(\bigvee_{j=0}^{n-1} U^j \mathscr{C}\Big) + H\Big(\bigvee_{i=0}^{n-1} U^i \mathscr{A} \Big| \bigvee_{j=0}^{n-1} U^j \mathscr{C}\Big).
\end{aligned}$$

Further $H(\mathscr{D} \vee \mathscr{E}|\mathscr{B}) \leqslant H(\mathscr{D}|\mathscr{B}) + H(\mathscr{E}|\mathscr{B}), H(\mathscr{D}|\mathscr{B} \vee \mathscr{E}) \leqslant H(\mathscr{D}|\mathscr{B})$ (see [5], [6]), hence

$$H\Big(\bigvee_{i=0}^{n-1} U^i\mathscr{A}\Big|\bigvee_{j=0}^{n-1} U^j\mathscr{C}\Big) \leqslant \sum_{i=0}^{n-1} H\Big(U^i\mathscr{A}\Big|\bigvee_{j=0}^{n-1} U^j\mathscr{C}\Big)$$
$$\leqslant \sum_{i=0}^{n-1} H(U^i\mathscr{A}|U^i\mathscr{C}) = nH(\mathscr{A}|\mathscr{C}).$$

Therefore

$$\lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} U^i\mathscr{A}\Big) \leqslant \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{j=0}^{n-1} U^j\mathscr{C}\Big) + H(\mathscr{A}|\mathscr{C}).$$

□

Lemma 4. *$h(U,\mathscr{C}) = h\Big(U, \bigvee_{j=0}^{k} U^j\mathscr{C}\Big)$ for every $k \in \mathbb{N}$ and any partition $\mathscr{C}$.*

P r o o f. We immediately obtain

$$\lim_{n\to\infty} \frac{1}{n}\Big(\bigvee_{i=0}^{n-1} U^i\Big(\bigvee_{j=0}^{k} U^j\mathscr{C}\Big)\Big) = \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{t=0}^{n+k-1} U^t\mathscr{C}\Big)$$
$$= \lim_{p\to\infty} \frac{p}{p-k} \cdot \frac{1}{p} H\Big(\bigvee_{t=0}^{p-1} U^t\mathscr{C}\Big) = h(U,\mathscr{C}).$$

□

P r o o f o f T h e o r e m. Let ε be an arbitrary positive number. Choose $\delta > 0$ by Lemma 2 and $\lambda > 0$ by Lemma 1. Now let $\mathscr{A} = \{f_1, \dots, f_n\} \subset \mathscr{G}$ be any partition. For every i there are $t_i \in \mathbb{N}$ and $g_i \in U^{t_i}\mathscr{C}$ such that $m(f_i \triangle g_i) < \lambda$ $(i = 1, 2, \dots, n)$. For sufficiently large k we have $\{g_1, \dots, g_n\} \subset \bigcup_{j=0}^{k} U^j\mathscr{C}$. By Lemma 1 there is a partition $\mathscr{B} = \{h_1, \dots, h_n\}$ such that $h_i \in \bigcup_{j=0}^{k} U^j\mathscr{C}$ and $m(f_i \triangle h_i) < \delta$ $(i = 1, \dots, n)$. By Lemma 2 we obtain $H(\mathscr{A}|\mathscr{B}) < \varepsilon$, by Lemma 4 we have

$$h\Big(U, \bigcup_{j=0}^{k} U^j\mathscr{C}\Big) = H(U,\mathscr{C}).$$

Therefore by Lemma 3 we conclude

$$\begin{aligned} h(U,\mathscr{A}) &\leqslant h(U,\mathscr{B}) + H(\mathscr{A}|\mathscr{B}) < h(U,\mathscr{B}) + \varepsilon \\ &\leqslant h\Bigl(U, \bigcup_{j=0}^{k} U^j\mathscr{C}\Bigr) + \varepsilon = h(U,\mathscr{C}) + \varepsilon, \\ H_{\mathscr{G}}(U) &= \sup\{H(U,\mathscr{A});\ \mathscr{A}\subset\mathscr{G}\} \leqslant h(U,\mathscr{C}). \end{aligned}$$

□

References

[1] *J. L. Bezdek*: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York, 1981.

[2] *P. Billingsley*: Ergodic Theory and Information. Willey, New York, 1965.

[3] *D. Butnariu, E. P. Klement*: Triangular norm-based measures and their Markov kernel representation. J. Math. Anal. Appl. *169* (1991), 111–143.

[4] *D. Dumitrescu*: Measure preserving transformation and the entropy of fuzzy partition. 13th Linz seminar on fuzzy set theory (Linz 1991), 25–27.

[5] *D. Dumitrescu*: Fuzzy measures and the entropy of fuzzy partitions. J. Math. Anal. Appl. *176* (1993), 359–373.

[6] *K. Kuriyama*: Entropy of a finite partition of fuzzy sets. J. Math. Anal. Appl. *94* (1983), 38–43.

[7] *P. Maličký, B. Riečan*: On the entropy of dynamical systems. Proc. Ergodic theory and related topics II (Georgenthal 1986). Teubner, Berlin, 1987, 135–138.

[8] *D. Markechová*: The entropy of F-quantum spaces. Math. Slovaca *40* (1990), 177–190.

[9] *D. Markechová*: The entropy of fuzzy dynamical systems and generators. Fuzzy Sets and Systems *48* (1992), 351–363.

[10] *R. Mesiar*: The Bayes formula and the entropy of fuzzy probability spaces. Int. J. General Systems *20* (1990), 67–71.

[11] *B. Riečan*: On a type of entropy of dynamical systems. Tatra Mountains Math. Publications.

[12] *D. J. Rudolph*: Fundamentals of Measurable Dynamics. Claredon Press, Oxford, 1990.

[13] *J. Rybárik*: The entropy of Q-F-dynamical systems. Busefal *48* (1991), 24–26.

Author's address: *Beloslav Riečan*, Matematický ústav SAV, Štefánikova 49, 814 73 Bratislava, Slovak Republic.