FUNDAMENTA
MATHEMATICAE
189 (2006)

Maličky–Riečan's entropy as a version of operator entropy

by

Bartosz Frej (Wrocław)

Abstract. The paper deals with the notion of entropy for doubly stochastic operators. It is shown that the entropy defined by Maličky and Riečan in [M-R] is equal to the operator entropy proposed in [D-F]. Moreover, some continuity properties of the [M-R] entropy are established.

1. Operator entropy. Let μ be a probability measure on a measurable space X. By a *doubly stochastic operator* we mean a linear operator T on $L^1(\mu)$ which satisfies the following conditions:

(i) Tf is positive for every positive $f \in L^1(\mu)$,
(ii) $T\mathbb{1} = \mathbb{1}$ (where $\mathbb{1}(x) = 1$ for all $x \in X$),
(iii) $\int Tf\, d\mu = \int f\, d\mu$ for every $f \in L^1(\mu)$.

It is well known that such an operator preserves each of the spaces $L^p(\mu)$ $(1 \le p \le \infty)$ and that a linear operator on $L^\infty(\mu)$ satisfying (i)–(iii) has a unique extension to a doubly stochastic operator on $L^1(\mu)$. Thus the name "doubly stochastic" may be used for operators on any $L^p(\mu)$ $(1 \le p \le \infty)$; the cases of $L^1(\mu)$, $L^2(\mu)$ and $L^\infty(\mu)$ are most often studied. Our considerations will not depend on the choice of the domain.

Since the Koopman operator of a measure-preserving transformation is doubly stochastic, one can interpret doubly stochastic operators as generalizations of classical dynamical systems. This justifies attempts to lift well known pointwise notions, like for instance entropy, to the operator case. Notice that the existence of a definition of entropy at the operator level does not collide with the well known fact that entropy in measure-preserving systems is not a spectral invariant. Several results have already been published in this area, the most recent being probably the theory developed in [D-F]. One of the theorems proved there states that all entropy notions set up for doubly stochastic operators are equal if only they follow some standard construction scheme (in particular, they are based on distinguished finite

2000 *Mathematics Subject Classification*: 28D20, 37A30, 37A35.
Key words and phrases: doubly stochastic operator, Markov operator, entropy.

collections of functions with an appropriate joining operation) and satisfy a few natural conditions like monotonicity, subadditivity, continuity (with respect to the collection of functions which replaces a partition) and some sort of compatibility with the classical entropy of a partition. Besides this axiomatic approach, one also needs a more explicit formula which would be convenient in proving properties of the entropy. To make the current paper self-contained we recall below the version of entropy proposed in [D-F] and formulate some of the results that can be found in that article.

For a function $f\colon X \to [0,1]$ let

$$A_f = \{(x,t) \in X \times [0,1] : t \le f(x)\}$$

and denote by $\mathcal{A}_f$ the partition of $X \times [0,1]$ consisting of A_f and its complement. For a collection $\mathcal{F}$ of measurable functions we define $\mathcal{A}_{\mathcal{F}} = \bigvee_{f\in\mathcal{F}} \mathcal{A}_f$. Clearly, $\mathcal{A}_{\mathcal{F}\cup\mathcal{G}} = \mathcal{A}_{\mathcal{F}} \vee \mathcal{A}_{\mathcal{G}}$. To shorten notation we write $T^n\mathcal{F} = \{T^n f : f \in \mathcal{F}\}$ and $\mathcal{F}^n = \bigcup_{k=0}^{n-1} T^k\mathcal{F}$, where $n \in \mathbb{N}$. Let λ be the Lebesgue measure on the unit interval. We define

$$\mathrm{H}^{\mathrm{DF}}(\mathcal{F}) = \mathrm{H}_{\mu\times\lambda}(\mathcal{A}_{\mathcal{F}}) = -\sum_{A\in\mathcal{A}_{\mathcal{F}}} (\mu\times\lambda)(A)\cdot\log(\mu\times\lambda)(A),$$

$$\mathrm{h}^{\mathrm{DF}}(T,\mathcal{F}) = \lim_{n\to\infty}\frac{1}{n}\mathrm{H}^{\mathrm{DF}}(\mathcal{F}^n), \qquad \mathrm{h}^{\mathrm{DF}}(T) = \sup_{\mathcal{F}} \mathrm{h}^{\mathrm{DF}}(T,\mathcal{F}),$$

with the supremum ranging over all finite collections of measurable functions from X to $[0,1]$. The existence of the limit in the second definition has been proved in [D-F]. The conditional entropy is given by

$$\mathrm{H}^{\mathrm{DF}}(\mathcal{F}\,|\,\mathcal{G}) = \mathrm{H}^{\mathrm{DF}}(\mathcal{F}\cup\mathcal{G}) - \mathrm{H}^{\mathrm{DF}}(\mathcal{G}).$$

Note that it is always nonnegative, since

$$\mathrm{H}^{\mathrm{DF}}(\mathcal{G}) \le \mathrm{H}^{\mathrm{DF}}(\mathcal{F}\cup\mathcal{G}) \le \mathrm{H}^{\mathrm{DF}}(\mathcal{F}) + \mathrm{H}^{\mathrm{DF}}(\mathcal{G}). \tag{1.1}$$

For a partition $\mathcal{A}$ of the underlying space X, let $\mathbb{1}_{\mathcal{A}}$ denote the set of characteristic functions of all elements of $\mathcal{A}$. The following property of H^{DF} is a concretization of the *domination axiom* from [D-F]:

For every $r \in \mathbb{N}$ and $\varepsilon > 0$ there exists a positive number γ such that for every collection $\mathcal{F} = \{f_1,\dots,f_r\}$ and every partition α of the unit interval into finitely many subintervals of lengths not exceeding γ we have

$$\mathrm{H}^{\mathrm{DF}}(\mathcal{F}\,|\,\mathbb{1}_{\bigvee_i f_i^{-1}(\alpha)} \cup \overline{\alpha}) < \varepsilon,$$

where $\overline{\alpha}$ is some set of constant functions, which depends only on α (not on $\mathcal{F}$).

For two collections of functions $\mathcal{F} = \{f_1,\dots,f_r\}$ and $\mathcal{G} = \{g_1,\dots,g_{r'}\}$, $r' \le r$, we define

$$\mathrm{dist}(\mathcal{F},\mathcal{G}) = \min_{\pi}\Big\{\max_{1\le i\le r}\int |f_i - g_{\pi(i)}|\,d\mu\Big\},$$

where the minimum ranges over all permutations π of the set $\{1,\ldots,r\}$ and where $\mathcal{G}$ is considered an r-element family by setting $g_i \equiv 0$ for $r' < i \le r$. For a function f and constants $a < b$ let $f_a^b = (f \vee a) \wedge b$, where $\vee$ and $\wedge$ denote maximum and minimum, respectively. We say that f has *property* $\mathrm{CZ}(\delta)$ if

$$\int |T^n(f_a^b) - (T^n f)_a^b|\, d\mu < \delta$$

for every $n \ge 0$ and every pair of constants $a < b$.

LEMMA 1.1 ([D-F, Lemma 2.3]). *For every bounded function f and every $\delta > 0$ there exists an integer l such that $T^l f$ has property* $\mathrm{CZ}(\delta)$.

LEMMA 1.2 ([D-F, Lemma 2.5]). *Let $\mathcal{F}$ consist of r functions with ranges in $[0,1]$, having property $\mathrm{CZ}(\delta^3)$. Then there is a partition Ξ of $[0,1]$ into subintervals of lengths not exceeding $2r\delta$ such that for every n and $f \in \mathcal{F}$,*

$$\int |T^n \mathbb{1}_{\{x:\, f(x) \ge \xi\}} - \mathbb{1}_{\{x:\, T^n f(x) \ge \xi\}}|\, d\mu < 4\delta,$$

where the values ξ are breakpoints of Ξ.

LEMMA 1.3 ([D-F, Lemma 2.6]). *Let $\mathcal{F} = \{f_1,\ldots,f_r\}$ consist of functions with ranges in $[0,1]$, all having property $\mathrm{CZ}(\delta^3)$, and let α be a partition of $[0,1]$ into m pieces $A_0 = [0,\xi_1)$, $A_j = [\xi_j,\xi_{j+1})$ $(j = 1,\ldots,m-2)$ and $A_{m-1} = [\xi_{m-1},1]$, where the points ξ_j all satisfy the assertion of Lemma* 1.2. *Then*

$$\mathrm{dist}(T^n(\mathbb{1}_{\bigvee_i f_i^{-1}(\alpha)}), \mathbb{1}_{\bigvee_i T^n f_i^{-1}(\alpha)}) < 8rm^r\delta$$

for every $n \ge 0$.

2. Maličky–Riečan's entropy. Another definition of entropy was proposed in [M-R] (a paper much earlier than [D-F]). It is based on the notion of a *partition of unity*, i.e. a finite collection of nonnegative, measurable functions whose sum is 1. For instance, if $\mathcal{A}$ is a finite measurable partition of the space X then $\mathbb{1}_{\mathcal{A}}$ forms a partition of unity. Instead of using some kind of joining operation, Maličky and Riečan introduce the following order relation on the set of partitions of unity: $\Psi \succeq \Phi$ if Ψ is the union of pairwise disjoint collections of functions Ψ_φ, $\varphi \in \Phi$, such that $\sum_{\psi \in \Psi_\varphi} \psi = \varphi$ (in fact, to ensure that the relation is antisymmetric we should exclude functions equal to 0 almost everywhere). Note that partitions of unity may have equal elements—for example $\{1/n, 1/n, \ldots, 1/n\}$ contains n elements although they are all identical.

The entropy of a partition of unity is given by

$$\mathrm{H}^{\mathrm{MR}}(\Phi) = -\sum_{\varphi \in \Phi} \int \varphi\, d\mu \cdot \log \int \varphi\, d\mu.$$

For several partitions $\Phi_1,\ldots,\Phi_n$ we put

$$\mathrm{H}^{\mathrm{MR}}(\Phi_1,\ldots,\Phi_n) = \inf\{\mathrm{H}^{\mathrm{MR}}(\Gamma) \colon \Gamma \succeq \Phi_1,\ldots,\Gamma \succeq \Phi_n\}.$$

Abbreviating $\mathrm{H}^{\mathrm{MR}}(\Phi, T\Phi, \ldots, T^{n-1}\Phi)$ by $\mathrm{H}^{\mathrm{MR}}(\Phi, n)$ we define

$$\mathrm{h}^{\mathrm{MR}}(T, \Phi) = \lim_{n\to\infty} \frac{1}{n} \mathrm{H}^{\mathrm{MR}}(\Phi, n).$$

It is not hard to verify that the sequence $\mathrm{H}^{\mathrm{MR}}(\Phi, n)$ is subadditive, thus the definition is correct. Finally, for any set $\mathcal{R}$ of partitions of unity we put

$$\mathrm{h}^{\mathrm{MR}}_{\mathcal{R}}(T) = \sup_{\Phi\in\mathcal{R}} \mathrm{h}^{\mathrm{MR}}(T, \Phi).$$

It is quite obvious that if T is a transformation and $\mathcal{R}$ contains all partitions of unity into measurable characteristic functions then $\mathrm{h}^{\mathrm{MR}}_{\mathcal{R}}(T)$ is not smaller than the classical Kolmogorov–Sinai entropy of T. The key tool here is the following easy lemma:

Lemma 2.1. *If $\Gamma \succeq \mathbb{1}_{\mathcal{A}}$ and $\Gamma \succeq \mathbb{1}_{\mathcal{A}'}$ for partitions $\mathcal{A}$, $\mathcal{A}'$ of the space X, then $\Gamma \succeq \mathbb{1}_{\mathcal{A}\vee\mathcal{A}'}$.*

We know from [M-R] that if T is a pointwise transformation and $\mathcal{R}$ is the set of all partitions of unity consisting of simple functions with rational coefficients then $\mathrm{h}^{\mathrm{MR}}_{\mathcal{R}}(T)$ is equal to the classical entropy of T. Because of the lack of a joining operation, Maličky–Riečan's construction does not follow the axiomatic scheme mentioned in the first section. Thus the equality between $\mathrm{h}^{\mathrm{MR}}_{\mathcal{R}}$ and h^{DF} for *any* doubly stochastic operator cannot be established by verification of the axioms. We devote the next section to proving this equality. Below we formulate a theorem which is a direct answer to the question posed in [M-R]—in the pointwise case, extending $\mathcal{R}$ from simple functions to all measurable partitions of unity does not increase $\mathrm{h}^{\mathrm{MR}}_{\mathcal{R}}$. Since this result is covered by Theorem 3.7, we refrain from presenting a detailed proof. However, we give a sketch of the argument, as it is based on some elementary calculations, without use of operator techniques.

Theorem 2.2. *If $T\colon X \to X$ is a measure preserving transformation and $\mathcal{R}$ denotes the set of all measurable partitions of unity then $\mathrm{h}^{\mathrm{MR}}_{\mathcal{R}}(T)$ is equal to the Kolmogorov–Sinai entropy of T.*

Sketch of proof. The goal is to find, for an arbitrary partition of unity $\Phi = \{\varphi_1, \ldots, \varphi_r\}$ and a positive number ε, a partition of unity $\mathcal{S}$ consisting of simple functions with rational coefficients and satisfying $\mathrm{h}^{\mathrm{MR}}(T, \Phi) \le \mathrm{h}^{\mathrm{MR}}(T, \mathcal{S}) + \varepsilon$. For $i = 1, \ldots, r$ we denote by s_i such a simple function approximating φ_i from below and define $\mathcal{S} = \{s_0, s_1, \ldots, s_r\}$, where $s_0 = 1 - \sum_{i=1}^{r} s_i$. Note that $\int s_0\, d\mu$ is small if the approximation is suitably exact. In addition we define a partition of unity $\mathcal{S}_\Phi$ consisting of $s_1, \ldots, s_r$ (common with $\mathcal{S}$) and new functions $s_{-i} = \varphi_i - s_i$ $(i = 1, \ldots, r)$.

Since $\mathcal{S}_\Phi \succeq \Phi$, we have $\mathrm{H}^{\mathrm{MR}}(\Phi, n) \le \mathrm{H}^{\mathrm{MR}}(\mathcal{S}_\Phi, n)$. We then show that the right hand side is dominated by $\mathrm{H}^{\mathrm{MR}}(\mathcal{S}, n) + (n+1)\varepsilon$. This is done in the following way: for a fixed n we pick a partition of unity Γ such that $\Gamma \succeq T^k\mathcal{S}$ for each $k < n$ and $\mathrm{H}^{\mathrm{MR}}(\Gamma) \le \mathrm{H}^{\mathrm{MR}}(\mathcal{S}, n) + \varepsilon$. For every $k < n$ the set Γ splits into disjoint subsets Γ_i^k whose elements sum to $T^k s_i$. For $\gamma \in \Gamma$ let $K(\gamma)$ denote the set of all numbers $k < n$ such that $\gamma \in \Gamma_0^k$, and let

$$P_i^k(x) = \begin{cases} \dfrac{T^k s_{-i}(x)}{T^k s_0(x)} & \text{if } T^k s_0(x) \neq 0, \\ 0 & \text{if } T^k s_0(x) = 0, \end{cases}$$

where i ranges over $\{1, \ldots, r\}$. If $K(\gamma)$ has $m > 0$ elements, we define for any m numbers $i_1, \ldots, i_m$ belonging to $\{1, \ldots, r\}$ the function

$$\widetilde{\gamma}_{i_1,\ldots,i_m} = \gamma \cdot \prod_{k \in K(\gamma)} P_{i_k}^k.$$

Clearly,

$$\sum_{i_1=1}^{r} \cdots \sum_{i_m=1}^{r} \widetilde{\gamma}_{i_1,\ldots,i_m} = \gamma.$$

Thus the set Γ_Φ consisting of all functions $\widetilde{\gamma}_{i_1,\ldots,i_m}$, $1 \le i_1, \ldots, i_m \le r$, and all elements $\gamma \in \Gamma$ for which $K(\gamma)$ is empty forms a partition of unity. Moreover, $\Gamma_\Phi \succeq T^k\mathcal{S}_\Phi$ for every $k < n$ and $|\mathrm{H}^{\mathrm{MR}}(\Gamma_\Phi) - \mathrm{H}^{\mathrm{MR}}(\Gamma)| < n\varepsilon$, since producing Γ_Φ we have modified a relatively small set of insignificant elements of Γ. □

3. Results. The definition of h^{DF} is based on partitions of the product $X \times [0,1]$, while $\mathrm{h}_{\mathcal{R}}^{\mathrm{MR}}$ exploits partitions of unity. In order to compare these notions we will introduce operations interchanging between partitions of unity and collections of functions inducing reasonable partitions of the product.

Let $\mathcal{F}$ be a collection of r functions $X \to [0,1]$. We may assume that the elements of $\mathcal{F}$ are pairwise unequal, because including in $\mathcal{F}$ multiple copies of the same function does not affect the partition $\mathcal{A}_{\mathcal{F}}$. We will say that $\mathcal{F}$ is *increasing* if its elements can be numbered in such a way that $f_{i+1} \ge f_i$ for $i = 1, \ldots, r-1$. From now on, every increasing collection of functions will be numbered in increasing order.

An arbitrary collection of functions may be transformed into an increasing one in the following way. Fix a numbering of $\mathcal{F} = \{f_1, \ldots, f_r\}$. Denote by "$\preccurlyeq$" the lexicographic order on the set of words $\{0,1\}^r$. For every $\beta \in \{0,1\}^r$ we define

$$\theta_\beta = \begin{cases} 1 & \text{for } \beta = 11\ldots1, \\ \sup\limits_{\alpha \preccurlyeq \beta} \inf\{f_i : \alpha_i = 0\} & \text{otherwise.} \end{cases}$$

It is clear that the functions θ_β form a collection which is increasing with respect to the lexicographic order on $\{0,1\}^r$—the outcome depends, however, on the initial numbering. Excluding from θ_β's spare copies of functions and adding the zero function we obtain an increasing collection which will be denoted by $\Theta(\mathcal{F})$. It seems intuitively obvious (and is not hard to prove) that

$$\mathcal{A}_{\Theta(\mathcal{F})} = \mathcal{A}_{\mathcal{F}}, \tag{3.1}$$

irrespective of the numbering of $\mathcal{F}$.

Next, we denote by $\mathcal{PU}(\mathcal{F})$ the partition of unity consisting of the functions $\theta_i - \theta_{i-1}$, where θ_i's are elements of $\Theta(\mathcal{F})$ in increasing order. Clearly,

$$\mathrm{H}^{\mathrm{MR}}(\mathcal{PU}(\mathcal{F})) = \mathrm{H}^{\mathrm{DF}}(\Theta(\mathcal{F})) = \mathrm{H}^{\mathrm{DF}}(\mathcal{F}). \tag{3.2}$$

On the other hand, an increasing collection of functions is obtained from a numbered partition of unity $\Phi = \{\varphi_1, \ldots, \varphi_s\}$ by means of the operation

$$\Sigma(\Phi) = \Big\{\sum_{i=1}^{j} \varphi_i \colon j = 0, 1, \ldots, s\Big\}$$

(the empty sum is interpreted as 0). Verification of the following easy formulas is left to the reader:

$$T\,\Sigma(\Phi) = \Sigma(T\Phi), \qquad \Sigma(\mathcal{PU}(\mathcal{F})) = \Theta(\mathcal{F}), \qquad \mathcal{PU}(\Sigma(\Phi)) = \Phi. \tag{3.3}$$

PROPOSITION 3.1. *Let $\mathcal{F} = \{f_0, \ldots, f_r\}$ and $\mathcal{G} = \{g_0, \ldots, g_s\}$ be increasing collections of functions with $f_0 = g_0 \equiv 0$, $f_r = g_s \equiv 1$. If $\mathcal{A}_{\mathcal{G}} \succcurlyeq \mathcal{A}_{\mathcal{F}}$ then $\mathcal{PU}(\mathcal{G}) \succeq \mathcal{PU}(\mathcal{F})$.*

Proof. The partition $\mathcal{A}_{\mathcal{F}}$ consists of the sets

$$A_i = \{(x, y) \colon f_{i-1}(x) < y \le f_i(x)\}, \qquad i = 1, \ldots, r.$$

For each i the set A_{i+1} lies "above" A_i in the sense that f_i is simultaneously an upper bound for A_i and a lower bound for A_{i+1}. Analogous statements are valid for $\mathcal{A}_{\mathcal{G}}$. Since every A_i is a union of elements of $\mathcal{A}_{\mathcal{G}}$, for every $i = 1, \ldots, r$ one can find integers $0 = j_0 < j_1 < \cdots < j_r = s$ such that $A_i = \bigcup_{j=j_{i-1}+1}^{j_i} B_j$, where $B_j = \{(x, y) \colon g_{j-1}(x) < y \le g_j(x)\}$. Thus for every $i = 1, \ldots, r$,

$$f_i - f_{i-1} = \sum_{j=j_{i-1}+1}^{j_i} (g_j - g_{j-1}),$$

i.e. $\mathcal{PU}(\mathcal{G}) \succeq \mathcal{PU}(\mathcal{F})$. ■

REMARK 3.2. The converse implication fails. For example, let $\mathcal{F}$ consist of the constant functions $1/3$ and 1, and let $\mathcal{G}$ consist of the constants $2/3$ and 1. The partitions $\mathcal{A}_{\mathcal{F}}$ and $\mathcal{A}_{\mathcal{G}}$ are not comparable, though $\mathcal{PU}(\mathcal{F})$ and $\mathcal{PU}(\mathcal{G})$ form the same partition of unity.

The next lemma is a part of the classical theory of entropy (cf. Proposition 11.10 of [D-G-S]).

LEMMA 3.3. *For every $\varepsilon > 0$ there exists $\delta > 0$ such that if $\mathcal{A} = \{A_1, \ldots, A_s\}$ and $\mathcal{A}' = \{A'_1, \ldots, A'_s\}$ are partitions of X satisfying $\mu(A_i \triangle A'_i) < \delta$ for $i = 1, \ldots, s$, then for any partition $\mathcal{B} \succcurlyeq \mathcal{A}$,*

$$\mathrm{H}_\mu(\mathcal{B} \vee \mathcal{A}') < \mathrm{H}_\mu(\mathcal{B}) + \varepsilon.$$

LEMMA 3.4. *For every $\varepsilon > 0$ and $s \in \mathbb{N}$ there exists $\delta > 0$ such that if Φ and Ψ are partitions of unity having at most s elements each and satisfying* $\operatorname{dist}(\Phi, \Psi) < \delta$, *then for any partition of unity $\Gamma \succeq \Phi$ there is a partition of unity $\widetilde{\Gamma}$ with the properties*:

(i) $\widetilde{\Gamma} \succeq \Gamma$,
(ii) $\widetilde{\Gamma} \succeq \Psi$,
(iii) $\mathrm{H}^{\mathrm{MR}}(\widetilde{\Gamma}) - \mathrm{H}^{\mathrm{MR}}(\Gamma) < \varepsilon$.

Proof. Number the elements of Φ and Ψ so that $\int |\varphi_i - \psi_i|\, d\mu < \delta$ (if the cardinalities of Φ and Ψ are not equal, we supply the smaller collection with an appropriate number of zeros). Then number Γ in such a way that the elements forming φ_i precede those used in the sum for φ_{i+1}. Define

$$\widetilde{\Gamma} = \mathcal{PU}(\Sigma(\Gamma) \cup \Sigma(\Psi)).$$

Then

$$\mathcal{A}_{\Sigma(\widetilde{\Gamma})} = \mathcal{A}_{\Theta(\Sigma(\Gamma) \cup \Sigma(\Psi))} \overset{(3.1)}{=} \mathcal{A}_{\Sigma(\Gamma) \cup \Sigma(\Psi)} = \mathcal{A}_{\Sigma(\Gamma)} \vee \mathcal{A}_{\Sigma(\Psi)}.$$

By Proposition 3.1 and (3.3), $\widetilde{\Gamma} \succeq \Gamma$ and $\widetilde{\Gamma} \succeq \Psi$.

Note that if $\operatorname{dist}(\Phi, \Psi) < \delta$ then $\operatorname{dist}(\Sigma(\Phi), \Sigma(\Psi)) < s\delta$ and the measure of the symmetric difference between each pair of corresponding elements of $\mathcal{A}_{\Sigma(\Phi)}$ and $\mathcal{A}_{\Sigma(\Psi)}$ is smaller than $2s\delta$. Since $\mathrm{H}^{\mathrm{MR}}(\mathcal{PU}(\mathcal{F})) = \mathrm{H}_{\mu \times \lambda}(\mathcal{A}_{\mathcal{F}})$, condition (iii) follows from the preceding lemma by letting $\mathcal{A} = \mathcal{A}_{\Sigma(\Phi)}$, $\mathcal{A}' = \mathcal{A}_{\Sigma(\Psi)}$, $\mathcal{B} = \mathcal{A}_{\Sigma(\Gamma)}$. ∎

LEMMA 3.5. *For every $\varepsilon > 0$ and $s \in \mathbb{N}$ there exists $\delta > 0$ such that if $\Phi_1, \ldots, \Phi_n$ and $\Psi_1, \ldots, \Psi_n$ are partitions of unity having at most s elements each and satisfying* $\operatorname{dist}(\Phi_i, \Psi_i) < \delta$ *for all $i = 1, \ldots, n$, then*

$$|\mathrm{H}^{\mathrm{MR}}(\Phi_1, \ldots, \Phi_n) - \mathrm{H}^{\mathrm{MR}}(\Psi_1, \ldots, \Psi_n)| < (n+1)\varepsilon.$$

Proof. Pick a partition of unity Γ_0 such that $\Gamma_0 \succeq \Phi_k$ for $k = 1, \ldots, n$ and $\mathrm{H}^{\mathrm{MR}}(\Gamma_0) \le \mathrm{H}^{\mathrm{MR}}(\Phi_1, \ldots, \Phi_n) + \varepsilon$. According to the previous lemma we choose a partition of unity Γ_1 satisfying $\Gamma_1 \succeq \Gamma_0$, $\Gamma_1 \succeq \Psi_1$ and $\mathrm{H}^{\mathrm{MR}}(\Gamma_1) - \mathrm{H}^{\mathrm{MR}}(\Gamma_0) < \varepsilon$. We continue the construction, in the kth step using the lemma for $\Phi = \Phi_k$, $\Psi = \Psi_k$ and $\Gamma = \Gamma_{k-1}$. After n steps we obtain a partition of unity Γ_n finer than Φ_k and Ψ_k for all $k = 1, \ldots, n$, and satisfying $\mathrm{H}^{\mathrm{MR}}(\Gamma_n) < \mathrm{H}^{\mathrm{MR}}(\Gamma_0) + n\varepsilon$. Hence

$$\mathrm{H}^{\mathrm{MR}}(\Psi_1, \ldots, \Psi_n) \le \mathrm{H}^{\mathrm{MR}}(\Gamma_n) < \mathrm{H}^{\mathrm{MR}}(\Phi_1, \ldots, \Phi_n) + (n+1)\varepsilon.$$

We finish the proof by exchanging the roles of Φ_k and Ψ_k in the above argument. ■

PROPOSITION 3.6 (Continuity of Maličky–Riečan's entropy). *The following continuity laws hold*:

(i) *For every $\varepsilon > 0$ and $s \in \mathbb{N}$ there exists $\delta > 0$ such that if Φ and Ψ have at most s elements each and satisfy* $\operatorname{dist}(\Phi, \Psi) < \delta$ *then for every $n \in \mathbb{N}$*,

$$|\mathrm{H}^{\mathrm{MR}}(\Phi, n) - \mathrm{H}^{\mathrm{MR}}(\Psi, n)| < n\varepsilon.$$

(ii) *For every $\varepsilon > 0$ and $s \in \mathbb{N}$ there exists $\delta > 0$ such that if Φ and Ψ have at most s elements each and satisfy* $\operatorname{dist}(\Phi, \Psi) < \delta$ *then*

$$|\mathrm{h}^{\mathrm{MR}}(T, \Phi) - \mathrm{h}^{\mathrm{MR}}(T, \Psi)| < \varepsilon.$$

Proof. The preceding lemma used for $\Phi_k = T^k\Phi$, $\Psi_k = T^k\Psi$ immediately implies (i), and (ii) follows from (i) and the definition of $\mathrm{h}^{\mathrm{MR}}(T, \Phi)$. ■

THEOREM 3.7. $\mathrm{h}^{\mathrm{MR}}_{\mathcal{R}}(T) = \mathrm{h}^{\mathrm{DF}}(T)$, *where $\mathcal{R}$ is the set of all measurable partitions of unity.*

Proof. We start by proving that $\mathrm{h}^{\mathrm{MR}}_{\mathcal{R}}(T) \le \mathrm{h}^{\mathrm{DF}}(T)$. Let Φ be a partition of unity. Denote by Φ^n_Σ the collection $\bigcup_{k=0}^{n-1} T^k(\Sigma(\Phi))$. For every $n \in \mathbb{N}$ and $k < n$ we have

$$\mathcal{A}_{\Theta(\Phi^n_\Sigma)} = \mathcal{A}_{\Phi^n_\Sigma} = \bigvee_{k=0}^{n-1} \mathcal{A}_{T^k(\Sigma(\Phi))} \succcurlyeq \mathcal{A}_{T^k(\Sigma(\Phi))}.$$

It follows from Lemma 3.1 that $\mathcal{PU}(\Phi^n_\Sigma) \succeq T^k\Phi$ for every $n \in \mathbb{N}$ and $k < n$. Thus

$$\mathrm{H}^{\mathrm{MR}}(\Phi, n) \le \mathrm{H}^{\mathrm{MR}}(\mathcal{PU}(\Phi^n_\Sigma)) \overset{(3.2)}{=} \mathrm{H}^{\mathrm{DF}}(\Phi^n_\Sigma).$$

Dividing both sides by n and letting $n \to \infty$ we get

$$\mathrm{h}^{\mathrm{MR}}(T, \Phi) \le \mathrm{h}^{\mathrm{DF}}(T, \Sigma(\Phi)).$$

For the reverse inequality we modify the proof of Theorem 2.1 of [D-F]. We fix $\varepsilon > 0$ and a collection $\mathcal{F} = \{f_1, \ldots, f_r\}$. Let γ be as specified in the domination axiom for the cardinality r and $\varepsilon > 0$. Choose m between $1/\gamma$ and $2/\gamma$. For $s = m^r$ and ε pick δ according to Lemma 3.5. Having in mind that $\mathrm{h}^{\mathrm{DF}}(T, \mathcal{F}) = \mathrm{h}^{\mathrm{DF}}(T, T^l\mathcal{F})$, we can use Lemma 1.1 to replace $\mathcal{F}$ by $T^l\mathcal{F}$ such that every $f \in \mathcal{F}$ has property $\mathrm{CZ}((\delta/8rm^r)^3)$. The number $\delta/4m^r$ majorizes the distances between the points ξ in Lemma 1.2 (with δ replaced by $\delta/8rm^r$), and because it is smaller than $\gamma/2$, we can pick $m-1$ of them creating a partition α of $[0,1]$ into m intervals of lengths smaller than γ. By the domination axiom, for every $k \in \mathbb{N}$ we have

$$\mathrm{H}^{\mathrm{DF}}(T^k\mathcal{F} \mid \mathbb{1}_{\bigvee_{f \in T^k\mathcal{F}} f^{-1}(\alpha)} \cup \overline{\alpha}) < \varepsilon,$$

so

$$\mathrm{H}^{\mathrm{DF}}(\mathcal{F}^n \mid \mathbb{1}_{\bigvee_{f\in\mathcal{F}^n} f^{-1}(\alpha)} \cup \overline{\alpha}) < n\varepsilon.$$

Lemma 1.3 yields

$$\operatorname{dist}(\mathbb{1}_{\bigvee_{f\in T^k\mathcal{F}} f^{-1}(\alpha)}, T^k(\mathbb{1}_{\bigvee_{f\in\mathcal{F}} f^{-1}(\alpha)})) < \delta. \tag{3.4}$$

Then

$$\begin{aligned}\mathrm{H}^{\mathrm{DF}}(\mathcal{F}^n) \quad &\le \quad \mathrm{H}^{\mathrm{DF}}(\mathbb{1}_{\bigvee_{f\in\mathcal{F}^n} f^{-1}(\alpha)} \cup \overline{\alpha}) + n\varepsilon\\ &\overset{(1.1),\ \text{Lem. } 2.1}{\le} \quad \mathrm{H}^{\mathrm{MR}}(\mathbb{1}_{\bigvee_{f\in\mathcal{F}} f^{-1}(\alpha)}, \ldots, \mathbb{1}_{\bigvee_{f\in T^{n-1}\mathcal{F}} f^{-1}(\alpha)})\\ &\qquad\quad + \mathrm{H}^{\mathrm{DF}}(\overline{\alpha}) + n\varepsilon.\end{aligned}$$

Using (3.4) and Lemma 3.5, we bound the right hand side of the above inequality by

$$\mathrm{H}^{\mathrm{MR}}(\mathbb{1}_{\bigvee_{f\in\mathcal{F}} f^{-1}(\alpha)}, n) + \mathrm{H}^{\mathrm{DF}}(\overline{\alpha}) + (2n+1)\varepsilon.$$

Thus $\mathrm{h}^{\mathrm{DF}}(T,\mathcal{F}) \le \mathrm{h}^{\mathrm{MR}}(T, \mathbb{1}_{\bigvee_{f\in\mathcal{F}} f^{-1}(\alpha)}) + 2\varepsilon \le \mathrm{h}^{\mathrm{MR}}_{\mathcal{R}}(T) + 2\varepsilon$. ∎

REMARK 3.8. One might hope to simplify the second part of the proof by extending Lemma 2.1 to collections of arbitrary functions. Unfortunately, it is not true that $\Gamma \succeq \Phi$, $\Gamma \succeq \Psi$ imply $\Gamma \succeq \mathcal{PU}(\Sigma(\Phi)\cup\Sigma(\Psi))$. For example, take $X = [0,1]$, $\Phi = \{x, 1-x\}$, $\Psi = \{\frac{1}{3}x + \frac{1}{3}, \frac{2}{3} - \frac{1}{3}x\}$ and $\Gamma = \{\gamma_1, \ldots, \gamma_6\}$, where $\gamma_1 = \gamma_2 = \gamma_3 = \frac{1}{3}x$, $\gamma_4 = \gamma_5 = \gamma_6 = \frac{1}{3} - \frac{1}{3}x$.

Acknowledgments. I am grateful to Wojciech Słomczyński for bringing the definition of Maličky and Riečan to my attention, and to Tomasz Downarowicz for fruitful discussions on the subject.

References

[D-G-S] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces*, Lecture Notes in Math. 527, Springer, Berlin, 1976.

[D-F] T. Downarowicz and B. Frej, *Measure-theoretic and topological entropy of operators on function spaces*, Ergodic Theory Dynam. Systems 25 (2005), 455–481.

[M-R] P. Maličky and B. Riečan, *On the entropy of dynamical systems*, in: Proc. Conf. Ergodic Theory and Related Topics II (Georgenthal, 1986), Teubner-Texte Math. 94, Teubner, Leipzig, 1987, 135–138.

Institute of Mathematics and Computer Science
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: frej@im.pwr.wroc.pl

Received 12 August 2005;
in revised form 20 September 2005