

Article

# On the Connections of Generalized Entropies With Shannon and Kolmogorov–Sinai Entropies

Fryderyk Falniowski

Department of Mathematics, Cracow University of Economics, Rakowicka 27, 31-510 Kraków, Poland;  
E-Mail: fryderyk.falniowski@uek.krakow.pl; Tel.: +48-122937500

Received: 17 May 2014; in revised form: 19 June 2014 / Accepted: 30 June 2014 /

Published: 3 July 2014

---

**Abstract:** We consider the concept of generalized Kolmogorov–Sinai entropy, where instead of the Shannon entropy function, we consider an arbitrary concave function defined on the unit interval, vanishing in the origin. Under mild assumptions on this function, we show that this isomorphism invariant is linearly dependent on the Kolmogorov–Sinai entropy.

**Keywords:** dynamical entropy; Kolmogorov–Sinai entropy

---

## 1. Introduction

The Boltzmann-Gibbs entropy and the Shannon entropy are measures that lie on the foundations of statistical physics and modern information theory, respectively. The dynamical counterparts of the latter—the dynamical and the Kolmogorov-Sinai entropy—had a great impact on the modern theory of dynamical systems and mathematical physics (see, e.g., [1,2]). They were extensively studied and successfully applied, among others, to statistical physics and quantum information. It appeared to be an exceptionally powerful tool for exploring nonlinear systems. One of the biggest advantages of the Kolmogorov-Sinai entropies lies in the fact that they make it possible to distinguish the formally regular systems (those with the measure-theoretic entropy equal to zero) from the chaotic ones (with positive entropy, which implies the positivity of topological entropy [3]).

The Kolmogorov-Sinai entropy of a given transformation  $T$  acting on a probability space  $(X, \Sigma, \mu)$  is defined as the supremum over all finite measurable partitions  $\mathcal{P}$  of the dynamical entropy of  $T$  with respect to  $\mathcal{P}$ , denoted by  $h(T, \mathcal{P})$ . As a dynamical counterpart of Shannon entropy, the entropy of

transformation  $T$  with respect to a given partition  $\mathcal{P}$  is defined as the limit of the sequence  $(\frac{1}{n}H(\mathcal{P}_n))_{n=1}^{\infty}$ , where:

$$H(\mathcal{P}_n) = \sum_{A \in \mathcal{P}_n} \eta(\mu(A))$$

with  $\eta$  being the Shannon function given by  $\eta(x) = -x \log x$  for  $x > 0$  with  $\eta(0) := 0$ , and  $\mathcal{P}_n$  is the join partition of partitions  $T^{-i}\mathcal{P}$  for  $i = 0, \dots, n - 1$ . The existence of the limit in the definition of the dynamical entropy follows from the subadditivity of  $\eta$ . The most common interpretation of this quantity is the average (over time and the phase space) one-step gain of information about the initial state. Taking the supremum over all finite partitions, we obtain an isomorphism invariant that measures the rate of producing randomness (chaos) by the system.

Since Shannon's seminal paper [4], several refinements of the concept of Shannon static entropy were considered (see, e.g., Rényi [5], Arimoto [6], Wu, Verdú [7] and Csiszár's survey article [8]). Their dynamical and measure-theoretic counterparts were also considered by a few authors. De Paly [9] proposed generalized dynamical entropies based on the concept of the relative static entropies. Unfortunately, it appeared that, despite some special cases [9,10], the explicit calculation of this invariant may not be possible. Grassberger and Procaccia proposed in [11] a dynamical counterpart of the well-known generalization of Shannon entropy—the Rényi entropy—and its measure-theoretic counterpart were considered by Takens and Verbitski. They showed that for ergodic transformations with positive measure-theoretic entropy, Rényi entropies of a measure-theoretic transformation are either infinite or equal to the measure-theoretic entropy [12]. The answer for non-ergodic aperiodic transformations is different: for Rényi entropies of order  $\alpha > 1$ , they are equal to the essential infimum of measure-theoretic entropies of measures forming the decomposition of a given measure into ergodic components, while for  $\alpha < 1$ , they are still infinite [13]. In particular, this means that Rényi entropies of order  $\alpha < 1$  are metric invariants sensitive to ergodicity. A similar generalization was made by Mesón and Vericat [14–16] for so-called Havrda–Charvát–Tsallis entropy [17], and their results were similar to the ones obtained by Takens and Verbitski in [12].

In this paper, we present a generalization of the concept of dynamical entropy. This was made for few reasons. First of all, the entropy used in the theory of dynamical systems is a natural isomorphism invariant. However, outside of the class of Bernoulli automorphisms, it is not a complete invariant, *i.e.*, two systems with the same entropy need not be isomorphic. In particular, it is not applicable to a wide class of systems with zero entropy. Moreover, we can ask whether considering of a function different than the Shannon function (at the expense of losing its interpretations from the information theory) can give a significantly new invariant. This can show which properties of the Shannon function are crucial for the usage of entropy in dynamical systems. Finally, different dynamical generalizations of entropy have been used by the physicists for years [18–22]. However, up till now, there have not been many rigorous results. We hope that this note will help to fill in this gap.

Our approach is based on the generalization suggested by Rényi and extended by Arimoto applied to the dynamical case. Instead of the Shannon function  $\eta$ , we consider a concave function  $g: [0, 1] \mapsto \mathbb{R}$ , such that  $\lim_{x \rightarrow 0^+} g(x) = g(0) = 0$  and define the dynamical  $g$ -entropy of the finite partition  $\mathcal{P}$  as:

$$h(g, T, \mathcal{P}) = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{A \in \mathcal{P}_n} g(\mu(A)).$$

The behavior of the quotient  $g(x)/\eta(x)$  as  $x$  converges to zero appears to be crucial for our considerations. Defining:

$$\text{Ci}(g) := \liminf_{x \rightarrow 0^+} \frac{g(x)}{\eta(x)} \quad \text{and} \quad \text{Cs}(g) := \limsup_{x \rightarrow 0^+} \frac{g(x)}{\eta(x)}$$

we will prove that:

$$\text{Ci}(g) \cdot h(T, \mathcal{P}) \leq h(g, T, \mathcal{P}) \leq \text{Cs}(g) \cdot h(T, \mathcal{P}).$$

In the case of  $\text{Ci}(g) = \infty$ , we will show that in every aperiodic system and for every  $\gamma \geq 0$ , there exists a finite partition  $\mathcal{P}$ , such that  $h(g, T, \mathcal{P}) \geq \gamma$ .

Considering the supremum over all partitions, we obtain a Kolmogorov entropy-like isomorphism invariant, which we will call the measure-theoretic  $g$ -entropy of a transformation with respect to an invariant measure. One might ask whether this invariant may give any new information about the system. We will prove (Theorem 5) that for  $g$  with  $\text{Cs}(g) < \infty$ , this new invariant is linearly dependent on Kolmogorov–Sinai entropy. It shows that Shannon entropy function is the special one from the point of view of the theory of dynamical entropies—it is the most natural one—not only does it have all of the properties that the entropy function should have [1], but also considering different entropy functions, we will not obtain an essentially different invariant. This result might have the other interpretation. Ornstein and Weiss showed in [23] that every finitely observable invariant for the class of all ergodic processes has to be a continuous function of the entropy. It is easy to see that any continuous function of the entropy is finitely observable; one simply composes the entropy estimators with the continuous function itself. In other words, an isomorphism invariant is finitely observable if and only if it is a continuous function of the Kolmogorov–Sinai entropy. Therefore, our result implies that the generalized measure-theoretic entropy is, in fact, finitely observable. It should be possible to give a more direct proof of the finite observability of the generalized measure-theoretic entropy, but the proof cannot be easier than the proof that entropy itself is finitely observable [24,25]. On the other hand, different entropy functions might be still of use, e.g., in the case of zero entropy systems, where we may consider generalizations of concepts used in this case: entropy convergence rates [26,27], generalized topological entropies [28] or entropy dimensions [29]. The generalization of entropy convergence rates can be found in [30]. The result implies that, from the point of view of the theory of entropy in dynamical systems, the crucial property of the Shannon function  $\eta$  is its behavior in the neighborhood of zero.

The note is organized as follows: in the next section, we introduce the dynamical  $g$ -entropy and establish its basic properties. The subsequent section is devoted to the construction of a zero

dynamical entropy process with sufficiently large  $g$ -entropy. Finally, in the last section, we define a measure-theoretic  $g$ -entropy of a transformation and show connections between this new invariant and the Kolmogorov-Sinai entropy.

## 2. Results

### 2.1. Basic Facts and Definitions

Let  $(X, \Sigma, \mu)$  be a Lebesgue space, and let  $g : [0, 1] \mapsto \mathbb{R}$  be a concave function with  $g(0) = \lim_{x \rightarrow 0^+} g(x) = 0$  (We might assume only that  $g(0) = 0$ , but then, the idea of the dynamical  $g$ -entropy would fail, since if  $\mathcal{P}_{n+1} \neq \mathcal{P}_n$  for every  $n$  and  $\lim_{x \rightarrow 0^+} g(x) \neq 0$ , then the dynamical  $g$ -entropy of the partition  $\mathcal{P}$  would be infinite. Therefore, if  $g$  is not well-defined at zero, we will assume that  $g(0) := \lim_{x \rightarrow 0^+} g(x)$ ). By  $\mathcal{G}_0$ , we will denote the set of all such functions. Every  $g \in \mathcal{G}_0$  is subadditive, i.e.,  $g(x + y) \leq g(x) + g(y)$  for every  $x, y, x + y \in [0, 1]$ , and quasi-homogenic, i.e.,  $\varphi_g : (0, 1] \rightarrow \mathbb{R}$  defined by  $\varphi_g(x) := g(x)/x$  is decreasing (see [31]). If  $g$  is fixed, we will omit the index, writing just  $\varphi$ .

Any finite family of pairwise disjoint subsets of  $X$ , such that  $\bigcup_{A_i \in \mathcal{P}} A_i = X$ , is called a partition. The set of all finite partitions will be denoted by  $\mathfrak{B}$ . For a given  $\mathcal{P} \in \mathfrak{B}$ , we define the  $g$ -entropy of the partition  $\mathcal{P}$  as:

$$H(g, \mathcal{P}) := \sum_{A \in \mathcal{P}} g(\mu(A)). \tag{1}$$

For  $g = \eta$ , the latter is equal to the Shannon entropy of the partition  $\mathcal{P}$ . For  $\mathcal{P}, \mathcal{Q} \in \mathfrak{B}$  of the space  $X$ , we define a new partition  $\mathcal{P} \vee \mathcal{Q}$  (the joint partition of  $\mathcal{P}$  and  $\mathcal{Q}$ ) consisting of the subsets of the form  $B \cap C$ , where  $B \in \mathcal{P}$  and  $C \in \mathcal{Q}$ .

#### 2.1.1. Dynamical $g$ -Entropies.

For an automorphism  $T : X \mapsto X$  and a partition  $\mathcal{P} = \{E_1, \dots, E_k\}$ , we put:

$$T^{-j}\mathcal{P} := \{T^{-j}E_1, \dots, T^{-j}E_k\}$$

and

$$\mathcal{P}_n = \mathcal{P} \vee T^{-1}\mathcal{P} \vee \dots \vee T^{-n+1}\mathcal{P}.$$

Now, for a given  $g \in \mathcal{G}_0$  and a finite partition  $\mathcal{P}$ , we can define the dynamical  $g$ -entropy of the transformation  $T$  with respect to  $\mathcal{P}$  as:

$$h_\mu(g, T, \mathcal{P}) = \limsup_{n \rightarrow \infty} \frac{1}{n} H(g, \mathcal{P}_n). \tag{2}$$

Alternatively we will call it the  $g$ -entropy of the process  $(X, \Sigma, \mu, T, \mathcal{P})$ . If the dynamical system  $(X, \Sigma, T, \mu)$  is fixed, then we omit  $T$ , writing just  $h(g, \mathcal{P})$ . As in the case of Shannon dynamical entropies, we are interested in the existence of the limit of  $(\frac{1}{n} H(g, \mathcal{P}_n))_{n=1}^\infty$ . If  $g = \eta$ , we obtain the

Shannon dynamical entropy  $h(T, \mathcal{P})$ . However, in the general case, we cannot replace an upper limit in Equation (2) by the limit, since it might not exist. The existence of the limit in the case of the Shannon function follows from the subadditivity of static Shannon entropy. This property has every subderivative function, *i.e.*, a function for which the inequality  $g(xy) \leq xg(y) + yg(x)$  holds for any  $x, y \in [0, 1]$ , but this is not true in general (an appropriate example will be given in Section 2.1.3). Therefore, we propose more general classes of functions for which this limit exists:

$$\mathcal{G}_0^0 := \left\{ g \in \mathcal{G}_0 \mid \lim_{x \rightarrow 0^+} \frac{g(x)}{\eta(x)} = 0 \right\} \quad \text{or} \quad \mathcal{G}_0^{\text{Sh}} := \left\{ g \in \mathcal{G}_0 \mid 0 < \lim_{x \rightarrow 0^+} \frac{g(x)}{\eta(x)} < \infty \right\}.$$

It is easy to show that if  $g$  is subderivative, then the limit  $\lim_{x \rightarrow 0^+} g(x)/\eta(x)$  is finite. Moreover, we will see that values of dynamical  $g$ -entropies depend on the behavior of  $g$  in the neighborhood of zero. We will prove that if  $g \in \mathcal{G}_0^0 \cup \mathcal{G}_0^{\text{Sh}}$ , then there is a linear dependence between the dynamical  $g$ -entropy and the Shannon dynamical entropy of a given partition. Before we give the general result (Theorem 1), we will state a few facts, which we will use in the proof of this theorem. We give the following lemmas, omitting their elementary proofs.

**Lemma 1.** *Let  $b_i > 0, a_i \in \mathbb{R}$  for  $i = 1, \dots, m$ , then*

$$\min_{i=1, \dots, m} \frac{a_i}{b_i} \leq \frac{\sum_{i=1}^m a_i}{\sum_{i=1}^m b_i} \leq \max_{i=1, \dots, m} \frac{a_i}{b_i}.$$

**Lemma 2.** *If  $\mathcal{P} \in \mathfrak{B}, \delta > 0$  and  $g: [0, 1] \mapsto \mathbb{R}$ , then:*

$$\sum_{A \in \mathcal{P}, \mu(A) \geq \delta} g(\mu(A)) \leq \frac{1}{\delta} \max_{x \in [\delta, 1]} g(x). \tag{3}$$

The following lemma states that the value of the dynamical  $g$ -entropy is given by the behavior of  $g$  in the neighborhood of zero.

**Lemma 3.** *If  $g_1, g_2 \in \mathcal{G}_0$  and there exists  $c > 0$ , such that  $g_1(x) = g_2(x)$  for  $x \in [0, c]$ , then for every  $\mathcal{P} \in \mathfrak{B}$   $h(g_1, \mathcal{P}) = h(g_2, \mathcal{P})$ .*

**Proof.** Let  $\mathcal{P} \in \mathfrak{B}$  and  $g_1, g_2 \in \mathcal{G}_0, c > 0$ , fulfill the assumptions. Because  $g \in \mathcal{G}_0$  is bounded, we have:

$$\begin{aligned} |H(g_1, \mathcal{P}_n) - H(g_2, \mathcal{P}_n)| &= \left| \sum_{A \in \mathcal{P}_n: \mu(A) > c} (g_1(\mu(A)) - g_2(\mu(A))) \right| \\ &\leq \frac{1}{c} \max_{x \in [c, 1]} |g_1(x) - g_2(x)|. \end{aligned}$$

Dividing by  $n$  and converging to infinity, we obtain:

$$h(g_1, \mathcal{P}) = h(g_2, \mathcal{P}).$$

□

We may state now the main theorem of this section.

**Theorem 1.** Let  $\mathcal{P} \in \mathfrak{B}$ .

- (1) If  $g \in \mathcal{G}_0$  is such that  $g'(0) < \infty$ , then  $h(g, \mathcal{P}) = 0$ .
- (2) If  $g_1, g_2 \in \mathcal{G}_0$  are such that  $g'_1(0) = g'_2(0) = \infty$ ,

$$\liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} < \infty,$$

and  $h(g_2, \mathcal{P}) < \infty$ , then:

$$\liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} \cdot h(g_2, \mathcal{P}) \leq h(g_1, \mathcal{P}).$$

If, additionally,  $\limsup_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} < \infty$ , then:

$$h(g_1, \mathcal{P}) \leq \limsup_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} \cdot h(g_2, \mathcal{P}).$$

- (3) If  $h(g_2, \mathcal{P}) = \infty$  and  $\liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} > 0$ , then  $h(g_1, \mathcal{P}) = \infty$ .

**Remark 1.** Whenever  $g_2: [0, 1] \mapsto \mathbb{R}$  is a nonnegative concave function satisfying  $g_2(0) = 0$  and  $g'_2(0) = \infty$ , we can have any pair  $0 < a \leq b \leq \infty$  as a limit inferior and limit superior of  $g_1/g_2$  in zero, choosing a suitable function  $g_1$ . The idea is as follows: construct  $g_1$  piecewise linear. To do so, define inductively a strictly decreasing sequence  $x_k \rightarrow 0$  and a decreasing sequence of values  $y_k = g_1(x_k) \rightarrow 0$ , thus defining intervals  $J_k := [x_{k+1}, x_k]$ , where  $g$  is affine. The only constraint to get a concave function is that the slope of  $g$  on each interval  $J_k$  has to be smaller than  $y_k/x_k$  and increasing with respect to  $k$ ; this is not an obstruction to approach any limit inferior and limit superior for  $g_1(x)/g_2(x)$ , provided that  $x_{k+1} > 0$  is chosen small enough.

**Proof of Theorem 1.** Let  $\mathcal{P} \in \mathfrak{B}$ . Suppose that  $g \in \mathcal{G}_0$  and  $g'(0) < \infty$ . Then:

$$h(g, \mathcal{P}) = \limsup_{n \rightarrow \infty} \frac{1}{n} H(g, \mathcal{P}_n) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \varphi(1/\text{card } \mathcal{P}_n) \leq \lim_{n \rightarrow \infty} \frac{g'(0)}{n} = 0,$$

which completes the proof of Point 1. To show Point 2, let  $g_1, g_2 \in \mathcal{G}_0$  be such that  $g'_1(0) = g'_2(0) = \infty$  and  $h(g_2, \mathcal{P}) < \infty$ . W.l.o.g, we can assume that  $g_1(x), g_2(x) > 0$  for  $x \in (0, 1)$ , since if there exists  $x_0 \in (0, 1)$ , such that  $g_i(x_0) = 0$  for  $i = 1$  or  $i = 2$ , then we can define  $\tilde{g}_i: [0, 1] \mapsto \mathbb{R}$  as:

$$\tilde{g}_i(x) := \begin{cases} g_i(x), & \text{for } x \in [0, s_i) \\ g_i(s_i), & \text{for } x \in [s_i, 1] \end{cases}$$

where  $s_i \in (0, 1]$  is such that  $\max_{x \in [0,1]} g(x) = g(s_i)$ . Then,  $\tilde{g}$  is strictly positive, and by Lemma 3 we have:

$$h(\tilde{g}_i, \mathcal{P}) = h(g_i, \mathcal{P}).$$

Let us assume that:

$$\limsup_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} < \infty.$$

Since  $g$  is subadditive, the sequence  $(H(g, \mathcal{P}_n))_{n=1}^\infty$  is nondecreasing, and there exists the limit of  $H(g_2, \mathcal{P}_n)$ . If it is finite, then  $h(g_2, \mathcal{P}) = 0$ , and by Equation (3) and Lemma 1, we have:

$$\begin{aligned} \sum_{A \in \mathcal{P}_n} g_1(\mu(A)) &\leq \sum_{A \in \mathcal{P}_n: \mu(A) < \frac{1}{2}} g_1(\mu(A)) + 2 \max_{x \in [\frac{1}{2}, 1]} g_1(x) \\ &\leq \sup_{x \in (0, \frac{1}{2})} \frac{g_1(x)}{g_2(x)} \cdot \sum_{A \in \mathcal{P}_n: \mu(A) < \frac{1}{2}} g_2(\mu(A)) + 2 \max_{x \in [\frac{1}{2}, 1]} g_1(x). \end{aligned}$$

Because  $\limsup_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} < \infty$ , there exists  $M > 0$ , such that  $g_1(x)/g_2(x) < M$  for  $x < 1/2$ . Therefore,

$\sup_{x \in (0, \frac{1}{2})} \frac{g_1(x)}{g_2(x)} < \infty$ , and by Lemma 1, we obtain:

$$\begin{aligned} 0 \leq h(g_1, \mathcal{P}) &= \limsup_{n \rightarrow \infty} \frac{1}{n} \frac{H(g_1, \mathcal{P}_n)}{H(g_2, \mathcal{P}_n)} H(g_2, \mathcal{P}_n) \\ &\leq \sup_{x \in (0, \frac{1}{2})} \frac{g_1(x)}{g_2(x)} \cdot \limsup_{n \rightarrow \infty} \frac{1}{n} H(g_2, \mathcal{P}_n) = 0. \end{aligned}$$

Thus, we can assume that  $\lim_{n \rightarrow \infty} H(g_2, \mathcal{P}_n) = \infty$ .

Fix  $\varepsilon > 0$ . There exists  $\delta > 0$ , such that, for  $x \in (0, \delta]$ , we have:

$$\liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} - \varepsilon < \frac{g_1(x)}{g_2(x)} \leq \limsup_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} + \varepsilon.$$

Lemma 1 implies that:

$$\liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} - \varepsilon \leq \frac{\sum_{A \in \mathcal{P}_n, \mu(A) < \delta} g_1(\mu(A))}{\sum_{A \in \mathcal{P}_n, \mu(A) < \delta} g_2(\mu(A))} \leq \limsup_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} + \varepsilon. \tag{4}$$

Using Equation (3) for every  $n > 0$ , we get:

$$\sum_{A \in \mathcal{P}_n, \mu(A) \geq \delta} g_i(\mu(A)) \leq \frac{1}{\delta} \overline{G}_\delta^i.$$

where  $\overline{G}_\delta^i := \max_{x \in [\delta, 1]} g_i(x)$  for  $i = 1, 2$ . Therefore:

$$\frac{\sum_{A \in \mathcal{P}_n: \mu(A) < \delta} g_1(\mu(A))}{\sum_{A \in \mathcal{P}_n: \mu(A) < \delta} g_2(\mu(A)) + \frac{1}{\delta} \overline{G}_\delta^2} \leq \frac{\sum_{A \in \mathcal{P}_n} g_1(\mu(A))}{\sum_{A \in \mathcal{P}_n} g_2(\mu(A))} \leq \frac{\sum_{A \in \mathcal{P}_n: \mu(A) < \delta} g_1(\mu(A)) + \frac{1}{\delta} \overline{G}_\delta^1}{\sum_{A \in \mathcal{P}_n: \mu(A) < \delta} g_2(\mu(A))}.$$

and  $\sum_{A \in \mathcal{P}_n: \mu(A) < \delta} g_2(\mu(A)) \rightarrow \infty (n \rightarrow \infty)$ . Dividing sums by  $\sum_{A \in \mathcal{P}_n: \mu(A) < \delta} g_2(\mu(A))$  and from Equation (4), we obtain:

$$\frac{\liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} - \varepsilon}{1 + \frac{\overline{G}_\delta^2}{\delta} \sum_{A \in \mathcal{P}_n: \mu(A) < \delta} g_2(\mu(A))} \leq \frac{\sum_{A \in \mathcal{P}_n} g_1(\mu(A))}{\sum_{A \in \mathcal{P}_n} g_2(\mu(A))} \leq \limsup_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} + \varepsilon + \frac{\overline{G}_\delta^1}{\delta} \sum_{A \in \mathcal{P}_n: \mu(A) < \delta} g_2(\mu(A))$$

Converging with  $n$  to infinity, we obtain:

$$\liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} - \varepsilon \leq \liminf_{n \rightarrow \infty} \frac{H(g_1, \mathcal{P}_n)}{H(g_2, \mathcal{P}_n)} \leq \limsup_{n \rightarrow \infty} \frac{H(g_1, \mathcal{P}_n)}{H(g_2, \mathcal{P}_n)} \leq \limsup_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} + \varepsilon.$$

Therefore:

$$\begin{aligned} \left( \liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} - \varepsilon \right) h(g_2, \mathcal{P}) &\leq \liminf_{n \rightarrow \infty} \frac{H(g_1, \mathcal{P}_n)}{H(g_2, \mathcal{P}_n)} \cdot \limsup_{n \rightarrow \infty} \frac{1}{n} H(g_2, \mathcal{P}_n) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} H(g_1, \mathcal{P}_n) \\ &\leq \limsup_{n \rightarrow \infty} \frac{H(g_1, \mathcal{P}_n)}{H(g_2, \mathcal{P}_n)} \cdot \limsup_{n \rightarrow \infty} \frac{1}{n} H(g_2, \mathcal{P}_n) \\ &\leq \left( \limsup_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} + \varepsilon \right) h(g_2, \mathcal{P}). \end{aligned}$$

Thus, we obtain the assertion. In the case of the infinite upper limit of the quotient  $g_1(x)/g_2(x)$ , we can repeat the above reasoning, just omitting the upper bound for the considered expressions.

If  $\liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} > 0$  and  $h(g_2, \mathcal{P}) = \infty$ , then  $\liminf_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} > \varepsilon$ , and using similar arguments, we obtain Point 3.  $\square$

Similar arguments lead us to the following statement:

**Theorem 2.** Let  $g_1, g_2 \in \mathcal{G}_0$  be such that  $\lim_{x \rightarrow 0^+} g_1(x)/g_2(x) = \infty$ , and let a finite partition  $\mathcal{P}$  have positive  $g_2$ -entropy. Then,  $h(g_1, \mathcal{P})$  is infinite.

Theorems 1 and 2 imply a few corollaries:

**Corollary 1.** If there exists the limit  $\lim_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} < \infty$ , then  $h(g_1, \mathcal{P}) = \lim_{x \rightarrow 0^+} \frac{g_1(x)}{g_2(x)} \cdot h(g_2, \mathcal{P})$ .

Let  $\mathcal{G}_0^\infty := \left\{ g \in \mathcal{G}_0 \mid \lim_{x \rightarrow 0^+} \frac{g(x)}{\eta(x)} = \infty \right\}$ . If  $g_1 = g, g_2 = \eta$ , then we have the following corollary:

**Corollary 2.** Let  $\mathcal{P} \in \mathfrak{B}$  and  $g \in \mathcal{G}_0$ , then:

(1) If  $\text{Ci}(g) < \infty$ , then  $h(g, \mathcal{P}) \geq \text{Ci}(g) \cdot h(\mathcal{P})$ .

(2) If  $Cs(g) < \infty$ , then  $h(g, \mathcal{P}) \in (Ci(g) \cdot h(\mathcal{P}), Cs(g) \cdot h(\mathcal{P}))$ .

(3) If  $g \in \mathcal{G}_0^0 \cup \mathcal{G}_0^{Sh}$ , then  $h(g, \mathcal{P}) = C(g) \cdot h(\mathcal{P})$ .

(4) If  $g \in \mathcal{G}_0^\infty$  and  $h(\mathcal{P}) > 0$ , then  $h(g, \mathcal{P}) = \infty$ .

**Corollary 3.** If  $(X, \Sigma, \mu, T)$  has positive Kolmogorov-Sinai entropy and  $g \in \mathcal{G}_0$  then:

$Cs(g) < \infty \Rightarrow g$ -entropy of any process  $(X, \Sigma, \mu, T, \mathcal{P})$  is finite  $\Rightarrow Ci(g) < \infty$ .

**Corollary 4.** If  $g \in \mathcal{G}_0^0 \cup \mathcal{G}_0^{Sh}$ , then  $h(g, \mathcal{P}) = \lim_{n \rightarrow \infty} \frac{1}{n} H(g, \mathcal{P}_n)$ .

### 2.1.2. Case of $g \in \mathcal{G}_0^\infty$

We will show that for every  $g \in \mathcal{G}_0^\infty$ , any aperiodic automorphism  $T$  and every  $\gamma \in \mathbb{R}$ , there exists a partition  $\mathcal{P} \in \mathfrak{B}$  such that  $h(g, \mathcal{P}) \geq \gamma$ . Since we omit the assumption of ergodicity, we will use different techniques, mainly based on the well-known Rokhlin Lemma, which guarantees the existence of so-called Rokhlin towers of a given height, covering a sufficiently large part of  $X$ . Using such towers, we will find lower bounds for the  $g$ -entropy of a process.

We will assume that we have an aperiodic system, i.e.,  $(X, \Sigma, \mu, T)$  for which

$$\mu(\{x \in X : \exists n \in \mathbb{N} T^n x = x\}) = 0.$$

If  $M_0, \dots, M_{n-1} \subset X$  are pairwise disjoint sets of equal measure, then  $\tau = (M_0, M_1, \dots, M_{n-1})$  is called a tower. If additionally  $M_k = T^{-(n-k-1)} M_{n-1}$  for  $k = 1, \dots, n - 1$ , then  $\tau$  is called a Rokhlin tower (It is also known as a Rokhlin–Halmos or Rokhlin–Kakutani tower.). By the same bold letter  $\tau$ , we will denote the set  $\bigcup_{k=0}^{n-1} M_k$ . Obviously  $\mu(\tau) = n\mu(M_{n-1})$ . Integer  $n$  is called the height of tower  $\tau$ . Moreover, for  $i < j$ , we define a sub-tower:

$$\tau_i^j := (M_i, \dots, M_j) \quad \text{and} \quad \tau_i^j = \bigcup_{k=i}^j M_k.$$

In aperiodic systems, there exist Rokhlin towers of a given length that cover a sufficiently large part of  $X$ :

**Lemma 4** ([32]). If  $T$  is an aperiodic and surjective transformation of Lebesgue space  $(X, \Sigma, \mu)$ , then for every  $\varepsilon > 0$  and every integer  $n \geq 2$ , there exists a Rokhlin tower  $\tau$  of height  $n$  with  $\mu(\tau) > 1 - \varepsilon$ .

Our goal is to find a lower bound for the dynamical  $g$ -entropy of a given partition. For this purpose, we will use Rokhlin towers, and we will calculate dynamical  $g$ -entropy with respect to a given Rokhlin tower. This leads to the following quantity: Let  $\mathcal{P}$  be a finite partition of  $X$  and  $F \in \Sigma$ ; then, we define the (static)  $g$ -entropy of  $\mathcal{P}$  restricted to  $F$  as:

$$H_F(g, \mathcal{P}) := \sum_{B \in \mathcal{P}} g(\mu(B \cap F)).$$

The following lemma gives the estimation for  $H(g, \mathcal{P})$  from below by the value of the  $g$ -entropy restricted to a subset of  $X$ .

**Lemma 5.** Let  $g \in \mathcal{G}_0$ . Let  $\mathcal{P}$  be a finite partition, such that there exists a set  $E \in \mathcal{P}$  with  $0 < \mu(E) < 1$ . If  $F \in \Sigma$ , then:

$$H(g, \mathcal{P}) \geq H_F(g, \mathcal{P}) - 3d_{\max},$$

where  $d_{\max} := \max_{x,y \in [0,1]} |g(x) - g(y)|$ .

**Proof.** Let  $A \in \mathcal{P}$ . Three-slope inequality implies that:

$$\frac{g(\mu(A)) - g(\mu(A \cap F))}{\mu(A) - \mu(A \cap F)} \geq \frac{g(1) - g(\mu(A))}{1 - \mu(A)}.$$

Thus, for sets of measure  $\mu(A) \leq \frac{1}{2}$ , we have:

$$\begin{aligned} g(\mu(A)) - g(\mu(A \cap F)) &\geq -\frac{\mu(A) - \mu(A \cap F)}{1 - \mu(A)} \cdot d_{\max} \\ &\geq -2(\mu(A) - \mu(A \cap F)) \cdot d_{\max}. \end{aligned}$$

Therefore, we obtain:

$$\sum_{A \in \mathcal{P}: \mu(A) \leq 1/2} (g(\mu(A)) - g(\mu(A \cap F))) \geq -2d_{\max} \cdot \sum_{A \in \mathcal{P}: \mu(A) \leq 1/2} \mu(A \setminus F) \geq -2d_{\max}$$

and:

$$\sum_{A \in \mathcal{P}: \mu(A) > 1/2} (g(\mu(A)) - g(\mu(A \cap F))) \geq -d_{\max},$$

which implies that:

$$H(g, \mathcal{P}) - H_F(g, \mathcal{P}) \geq -3d_{\max}.$$

□

The following lemma will play an important role in the proof of the main theorem of this section.

**Lemma 6.** Let  $n \in \mathbb{N}$ ,  $E \in \Sigma$ . Let  $\mathcal{P}^E := \{E, X \setminus E\}$ . Suppose that  $g \in \mathcal{G}_0$  is nonnegative in  $[0, \alpha]$ , where  $\alpha$  is some positive number. Then, there exist  $\delta > 0$  and  $s \in (0, \alpha)$ , such that:

$$|H(g, \mathcal{P}_n^E) - H(g, \mathcal{P}_n^F)| \leq 1 + \frac{2}{s}d_{\max}$$

for every  $F \in \Sigma$  s.t.  $\mu(E \Delta F) < \delta$  ( $\Delta$  denotes the symmetric difference), where  $d_{\max} := \max_{x,y \in [0,1]} |g(x) - g(y)|$ .

**Proof.** It is easy to show that for every  $n \in \mathbb{N}$ ,  $E \in \Sigma$ ,  $\varepsilon > 0$ , there exists  $\delta > 0$ , such that:

$$\min_{\pi} \sum_{A_i \in \mathcal{P}_n^E, B_{\pi(i)} \in \mathcal{P}_n^F} \mu(A_i \Delta B_{\pi(i)}) < \varepsilon \text{ for } F \in \Sigma \text{ such that: } \mu(E \Delta F) < \delta.$$

W.l.o.g., we may assume that  $\sum_{i=1}^{2^n} \mu(A_i \Delta B_i) < \varepsilon$  (adding empty sets if necessary). The nonnegativity of  $g$  for  $x \in [0, \alpha]$  and its concavity imply that there exists  $s \in (0, \alpha)$ , such that  $g$  is nondecreasing in  $[0, s]$ . Fix  $n \in \mathbb{N}$  and  $E \in \Sigma$ . There exists  $\varepsilon \in (0, s/2)$ , such that:

$$g(\varepsilon) < 2^{-n}. \tag{5}$$

It is easy to see that for  $x \in [0, s]$ , the monotonicity and subadditivity of  $g$  implies that:

$$|g(y) - g(x)| \leq g(|y - x|). \tag{6}$$

Let  $F \in \Sigma$  be such that  $\mu(E \Delta F) \leq \delta$ . Define  $\mathcal{D}_s^n = \{i \in \{1, \dots, 2^n\} \mid \max\{\mu(A_i), \mu(B_i)\} < s\}$ . From Equations (5) and (6) and the monotonicity of  $g$  in  $[0, s]$ , we obtain:

$$\begin{aligned} |H(g, \mathcal{P}_n^E) - H(g, \mathcal{P}_n^F)| &\leq \sum_{i \in \mathcal{D}_s^n} |g(\mu(A_i)) - g(\mu(B_i))| + \frac{2}{s} d_{\max} \\ &\leq \sum_{i \in \mathcal{D}_s^n} g(|\mu(A_i) - \mu(B_i)|) + \frac{2}{s} d_{\max} \\ &\leq \sum_{i \in \mathcal{D}_s^n} g(\mu(A_i \Delta B_i)) + \frac{2}{s} d_{\max} \\ &\leq \sum_{i=1}^m g(\mu(A_i \Delta B_i)) + \frac{2}{s} d_{\max} \\ &\leq 2^n g(\delta) + \frac{2}{s} d_{\max} \leq 1 + \frac{2}{s} d_{\max} \end{aligned}$$

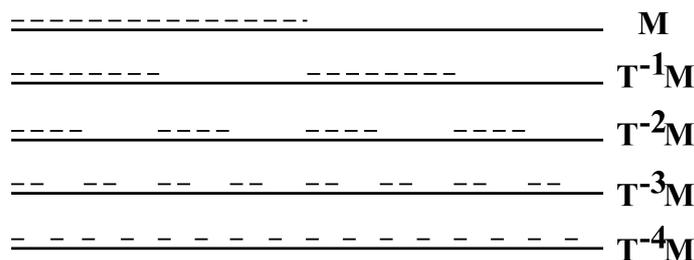
□

To find the lower bound for the  $g$ -entropy of a partition, we will use so-called independent sets. We construct the independent set in the following way: Let  $\tau$  be a tower of height  $m$ . We divide the highest level of this tower ( $M_{m-1}$ ) into two sets of equal measure, let us say  $I^{(m-1)}$  and  $M_{m-1} \setminus I^{(m-1)}$ . Next, we consider  $T^{-1}I^{(m-1)}$  and  $T^{-1}(M_{m-1} \setminus I^{(m-1)})$ . We divide each of them into two sets of equal measure, obtaining sets  $I_1^{(m-2)}, I_2^{(m-2)}, I_3^{(m-2)}, I_4^{(m-2)}$ , and define set  $I^{(m-2)}$  as the algebraic sum of two of those sets—one subset of  $T^{-1}I^{(m-1)}$  and one of  $T^{-1}(M_{m-1} \setminus I^{(m-1)})$ . We perform this algorithm, until we achieve the lowest level of the tower —  $M_0$  (see Figure 1). Eventually, we define  $I := \bigcup_{j=0}^{m-1} I^{(j)}$ . We call this set an independent set in  $\tau$ .

We can make this construction, because every aperiodic system does not have atoms of positive measure, and in every non-atomic Lebesgue space for every measurable set  $A$  and every  $\alpha \in [0, \alpha]$ , there exists  $B \subset A$ , such that  $\mu(B) = \alpha$ .

We are able to give an explicit formula for the  $g$ -entropy of the partition generated by the independent set in  $\tau$ .

**Figure 1.** Set  $I$  (with dashes) in a tower a height of five.



**Lemma 7.** Let  $\tau = (M, TM, \dots, T^{2n-1}M)$  be the Rokhlin tower of height  $2n$  and  $I \in \Sigma$  be an independent set in  $\tau$ . If  $g \in \mathcal{G}_0^\infty$ , then:

$$H_{\tau_0^{n-1}}(g, \mathcal{P}_n^I) = \frac{\mu(\tau)}{2} \varphi\left(\frac{\mu(\tau)}{2^{n+1}}\right),$$

where  $\varphi(x) = g(x)/x$  for  $x > 0$ .

**Proof.** The independence of  $I$  in  $\tau$  implies that the partition:

$$\mathcal{P}_n^I \cap \tau_0^{n-1}$$

is a partition of  $\tau_0^{n-1}$  into  $2^n$  sets of equal measure  $2^{-n}\mu(\tau_0^{n-1})$ . Therefore:

$$\begin{aligned} H_{\tau_0^{n-1}}(g, \mathcal{P}_n^I) &= \sum_{A \in \mathcal{P}_n^I} g(\mu(A \cap \tau_0^{n-1})) \\ &= 2^n g\left(\frac{\mu(\tau_0^{n-1})}{2^n}\right) = \mu(\tau_0^{n-1}) \varphi\left(\frac{\mu(\tau_0^{n-1})}{2^n}\right) \\ &= \frac{\mu(\tau)}{2} \varphi\left(\frac{\mu(\tau)}{2^{n+1}}\right). \end{aligned}$$

□

**Theorem 3.** Let  $g \in \mathcal{G}_0^\infty$  and  $T$  be an aperiodic, surjective automorphism of a Lebesgue space  $(X, \Sigma, \mu)$ , and let  $\gamma \in \mathbb{R}$ , then, there exists a partition  $\mathcal{P} \in \mathfrak{B}$ , such that:

$$h(g, \mathcal{P}) \geq \gamma.$$

**Proof.** We will prove that for any  $\gamma > 0$ , there exists a partition  $\mathcal{P}^E = \{E, X \setminus E\}$ , such that  $h(g, \mathcal{P}) \geq \gamma$ . We define recursively a sequence of sets  $E_n \in \Sigma$ . Let:

$$E_0 := \emptyset, \quad N_0 := \delta_0 := 1.$$

Let  $n > 0$ , and assume that we have already defined  $E_{n-1}, N_{n-1}$  and  $\delta_{n-1}$ . Using Lemma 6, we can choose  $\delta_n > 0$ , such that:

$$\delta_n < \frac{1}{2} \delta_{n-1} \tag{7}$$

$$\left| H(g, \mathcal{P}_{N_n}^{E_{n-1}}) - H(g, \mathcal{P}_{N_n}^F) \right| < 1 + \frac{2}{s} d_{\max}. \tag{8}$$

for any  $F \in \Sigma$ , for which  $\mu(E_{n-1} \Delta F) < 2\delta_n$ .

Since:

$$\lim_{x \rightarrow 0^+} \frac{g(x)}{\eta(x)} = \infty,$$

we can choose such  $N_n \in \mathbb{N}$  that:

$$\frac{\varphi(\delta_n 2^{-N_n-1})}{\varphi_\eta(\delta_n 2^{-N_n-1})} > \frac{2\gamma}{\delta_n \log 2}. \tag{9}$$

By Lemma 4, there exists  $M_n \in \Sigma$ , such that  $\tau_n = (M_n, TM_n, \dots, T^{2^{N_n-1}}M_n)$  is a Rokhlin tower of measure  $\mu(\tau_n) = \delta_n$ . Let  $I_n \subset \tau_n$  be an independent set in  $\tau_n$  and:

$$E_n := (E_{n-1} \setminus \tau_n) \cup I_n.$$

Then:

$$\mu(E_{n-1} \Delta E_n) \leq \mu(\tau_n) = \delta_n.$$

for all positive integers  $n$ . By Equation (7), we have  $\delta_n < 2^{-n}$ , and we conclude that  $(\mathbf{1}_{E_n})_{n=0}^\infty$  is a Cauchy sequence in  $L_1(X)$ . Therefore, there exist  $E \in \Sigma$ , such that  $\mathbf{1}_{E_n}$  converges to  $\mathbf{1}_E$ . For this set, we have:

$$\mu(E_n \Delta E) \leq \sum_{k=n+1}^\infty \mu(E_k \Delta E_{k-1}) \leq \sum_{k=n+1}^\infty \delta_k < 2\delta_{n+1}.$$

Since  $E_n \cap \tau_n = I_n$ , applying Equation (8) and Lemmas 5 and 7, we obtain that for  $N_n$ , such that  $\delta_n \cdot 2^{-N_n-1} < s$ :

$$\begin{aligned} H(g, \mathcal{P}_{N_n}^E) &\geq H(g, \mathcal{P}_{N_n}^{E_n}) - 1 - \frac{2}{s} d_{\max} \\ &\geq H_{(\tau_n)_0^{N_n-1}}(g, \mathcal{P}_{N_n}^{E_n}) - \left(\frac{2}{s} + 3\right) d_{\max} - 1 \\ &\geq H_{(\tau_n)_0^{N_n-1}}(g, \mathcal{P}_{N_n}^{I_n}) - \left(\frac{2}{s} + 3\right) d_{\max} - 1 \\ &\geq \left[ \frac{\mu(\tau_n) \ln 2}{2} (N_n + 1) - \frac{\mu(\tau_n) \ln \mu(\tau_n)}{2} \right] \cdot \frac{\varphi(\mu(\tau_n) 2^{-N_n-1})}{-\ln(\mu(\tau_n) 2^{-N_n-1})} \\ &\quad - \left(\frac{2}{s} + 3\right) d_{\max} - 1 \\ &\geq \frac{\ln 2}{2} \cdot \delta_n \cdot (N_n + 1) \cdot \frac{\varphi(\delta_n 2^{-N_n-1})}{\varphi_\eta(\delta_n 2^{-N_n-1})} - \left(\frac{2}{s} + 3\right) d_{\max} - 1. \end{aligned}$$

From Equation (9), we obtain that:

$$\lim_{n \rightarrow \infty} \frac{H(g, \mathcal{P}_{N_n}^E)}{N_n} \geq \frac{\ln 2}{2} \lim_{n \rightarrow \infty} \delta_n \cdot \frac{N_n + 1}{N_n} \cdot \frac{\varphi(\delta_n 2^{-N_n-1})}{\varphi_\eta(\delta_n 2^{-N_n-1})} \geq \gamma.$$

Thus,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} H(g, \mathcal{P}_n) \geq \gamma.$$

□

### 2.1.3. Bernoulli Shifts

Let  $\mathcal{A} = \{1, \dots, k\}$  be a finite alphabet. Let  $X = \{x = \{x_i\}_{i=-\infty}^{\infty} : x_i \in \mathcal{A}\}$  and  $\sigma$  be a left shift:

$$\sigma(x)_i = x_{i+1}.$$

For any  $s \leq t$  and block  $[\omega_0, \dots, \omega_{t-s}]$  with  $a_i \in \mathcal{A}$ , we define a cylinder:

$$C_s^t(\omega_0, \dots, \omega_{t-s}) = \{x \in X : x_i = \omega_{i-s} \text{ for } i = s, \dots, t\}.$$

We consider a Borel  $\sigma$ -algebra with respect to the metric, which is given by  $d(x, y) = 2^{-N}$ , where  $N = \min\{|i| : x_i \neq y_i\}$ . Let  $p = (p_1, \dots, p_k)$  be a probability vector. We define a measure  $\rho = \rho(p)$  on  $\mathcal{A}$  by setting  $\rho(\{i\}) = p_i$ . Then,  $\mu_p$  is a corresponding product measure on  $X = \mathcal{A}^{\mathbb{Z}}$ . Thus, the static  $g$ -entropy of a partition  $\mathcal{P}^{\mathcal{A}} = \{[1], [2], \dots, [k]\}$  is equal to:

$$H_{\mu_p}(g, \mathcal{P}_n^{\mathcal{A}}) = \sum_{\omega \in \mathcal{A}^n} g(\mu(C_0^{n-1}(\omega_0, \dots, \omega_{n-1}))) = \sum_{\omega \in \mathcal{A}^n} g(p_{\omega_0} \cdots p_{\omega_{n-1}}),$$

where  $\omega = (\omega_0, \dots, \omega_{n-1})$ . By the concavity of  $g$ , we have:

$$H_{\mu_p}(g, \mathcal{P}_n^{\mathcal{A}}) \leq \varphi(1/k^n)$$

where equality holds only when  $p = p^* = (\frac{1}{k}, \dots, \frac{1}{k})$ . Before calculating the dynamical  $g$ -entropy of  $\mathcal{P}^{\mathcal{A}}$  with respect to  $\mu_{p^*}$ , we give the following lemma, the proof of which will be given later:

**Lemma 8.** *If  $g \in \mathcal{G}_0$ , then:*

$$Cs(g) = \limsup_{n \rightarrow \infty} \frac{g(\kappa^{-n})}{\eta(\kappa^{-n})} \text{ and } Ci(g) = \liminf_{n \rightarrow \infty} \frac{g(\kappa^{-n})}{\eta(\kappa^{-n})}$$

for any  $\kappa > 1$ .

Therefore, applying Lemma 8 for the partition  $\mathcal{P}^{\mathcal{A}}$  and  $\kappa = k$ , we obtain:

$$h_{\mu_{p^*}}(g, \mathcal{P}^{\mathcal{A}}) = \limsup_{n \rightarrow \infty} \frac{1}{n} \varphi\left(\frac{1}{k^n}\right) = \begin{cases} Cs(g) \cdot \log k, & \text{if } Cs(g) < \infty; \\ \infty, & \text{otherwise.} \end{cases} \tag{10}$$

**Remark 2.** *If we consider the lower limit instead of the upper limit, we would obtain:*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \varphi\left(\frac{1}{k^n}\right) = \begin{cases} Ci(g) \cdot \log k, & \text{if } Ci(g) < \infty; \\ \infty, & \text{otherwise.} \end{cases}$$

Therefore, we cannot replace an upper limit by the limit in the definition of the dynamical  $g$ -entropy.

**Proof of Lemma 8.** We will show just the equality for the upper limit, since the equality for the lower limit may be obtained analogously. Let  $(x_n)_{n=1}^\infty$  and  $(m_n)_{n=1}^\infty$  be such that  $\limsup_{n \rightarrow \infty} g(x_n)/\eta(x_n) = c$  and  $x_n \in (\kappa^{-m_n}, \kappa^{-m_n+1})$  for every  $n \in \mathbb{N}$ . Then,  $-\log x_n \geq -\log \kappa^{-m_n+1}$ . Every  $g \in \mathcal{G}_0$  is quasi-homogenic; so, for every positive integer  $n$  occurs:

$$\frac{g(x_n)}{x_n} < \frac{g(\kappa^{-m_n})}{\kappa^{-m_n}}.$$

Therefore:

$$\begin{aligned} \frac{g(x_n)}{\eta(x_n)} &= \frac{g(x_n)}{x_n} \frac{1}{-\log x_n} \leq \frac{g(\kappa^{-m_n})}{\kappa^{-m_n}} \frac{1}{(m_n - 1) \log \kappa} \\ &= \frac{g(\kappa^{-m_n})}{\eta(\kappa^{-m_n})} \cdot \frac{m_n}{m_n - 1}, \end{aligned}$$

and:

$$\limsup_{x \rightarrow 0^+} \frac{g(x)}{\eta(x)} = \limsup_{n \rightarrow \infty} \frac{g(\kappa^{-n})}{\eta(\kappa^{-n})}.$$

□

## 2.2. Kolmogorov-Sinai Entropy-Like Invariant

The basic tool considered in the ergodic theory is the Kolmogorov–Sinai entropy, which is a supremum of Shannon dynamical entropies over all finite partitions:

$$h_\mu(T) = \sup_{\mathcal{P} \in \mathfrak{B}} h(T, \mathcal{P}).$$

It is invariant under metric isomorphism. Following the Kolmogorov proposition, we take the supremum over all partitions of the dynamical  $g$ -entropy of a partition. For a given system  $(X, \Sigma, \mu, T)$ , we define:

$$h_\mu(g, T) = \sup_{\mathcal{P} \in \mathfrak{B}} h(g, T, \mathcal{P}) \tag{11}$$

and call it the measure-theoretic  $g$ -entropy of transformation  $T$  with respect to measure  $\mu$ .

It is easy to see that it is an isomorphism invariant. Ornstein and Weiss [23] showed the striking result that measure-theoretic entropy is the only finitely observable invariant for the class of all ergodic processes. More precisely, every finitely observable invariant for a class of all ergodic processes is a continuous function of entropy. Of course, in the case of  $g \in \mathcal{G}_0^0 \cup \mathcal{G}_0^{\text{Sh}}$  by Corollary 2, we have:

$$h_\mu(g, T) = \lim_{x \rightarrow 0^+} \frac{g(x)}{\eta(x)} \cdot h_\mu(T).$$

We will show that for a wider class of functions, namely for those for which:

$$Cs(g) = \limsup_{x \rightarrow 0^+} \frac{g(x)}{\eta(x)} < \infty$$

we have:

$$h_\mu(g, T) = Cs(g) \cdot h_\mu(T)$$

for any ergodic transformation  $T$ . This shows that the measure-theoretic  $g$ -entropy is, in fact, finitely observable: one might simply compose the entropy estimators [25] with the linear function itself. Our proof will be similar to the proof of ([12], Theorem 1.1) where Takens and Verbitski showed that for ergodic transformations, the supremum over all finite partitions of dynamical Rényi entropies of order  $\alpha > 1$  are equal to the measure-theoretic entropy of  $T$  with respect to measure  $\mu$ .

Let us introduce necessary definitions and facts. Let  $T_i$  be automorphisms of Lebesgue space  $(X_i, \Sigma_i, \mu_i)$  for  $i = 1, 2$ , respectively. Then, we say that  $T_2$  is a factor of transformation  $T_1$ , if there exists a homomorphism  $\phi: X_1 \mapsto X_2$ , such that:

$$\phi T_1 = T_2 \phi \quad \mu_1 \text{ a.e. on } X_1.$$

Suppose that  $T_2$  is a factor of  $T_1$  under homomorphism  $\phi$ . Then, for an arbitrary finite partition  $\mathcal{P}$  of  $X_2$ , we have:

$$H \left( g, \bigvee_{i=0}^{k-1} T_2^{-i} \mathcal{P} \right) = H \left( g, \bigvee_{i=0}^{k-1} \phi^{-1} T_2^{-i} \mathcal{P} \right) = H \left( g, \bigvee_{i=0}^{k-1} T_1^{-i} \phi^{-1} \mathcal{P} \right).$$

Hence,  $h(g, T_2, \mathcal{P}) = h(g, T_1, \phi^{-1} \mathcal{P})$ . Therefore:

$$h_\mu(g, T_2) = \sup_{\mathcal{P}\text{-finite}} h(g, T_2, \mathcal{P}) = \sup_{\mathcal{P}\text{-finite}} h(g, T_1, \phi^{-1} \mathcal{P}) \leq h(g, T_1).$$

This implies the following proposition:

**Proposition 1.** *If  $T_2$  is a factor of  $T_1$ , then for every function  $g \in \mathcal{G}_0$ :*

$$h_\mu(g, T_2) \leq h_\mu(g, T_1).$$

### 2.2.1. Measure-Theoretic $g$ -Entropies for Bernoulli Automorphisms

An automorphism  $T$  on  $(X, \Sigma, \mu)$  is called Bernoulli automorphism if it is isomorphic to some Bernoulli shift. The crucial role in the proof of the main theorem of this section (Theorem 5) will be played by a well-known Sinai theorem:

**Theorem 4** (Sinai [33]). *Let  $T$  be an arbitrary ergodic automorphism of some Lebesgue space  $(X, \Sigma, \mu)$ . Then, each Bernoulli automorphism with  $h_\mu(T_1) \leq h_\mu(T)$  is a factor of the automorphism  $T$ .*

We start proving the following proposition:

**Proposition 2.** Let  $T$  be an arbitrary ergodic automorphism with  $h_\mu(T) \geq \log M$  for some integer  $M \geq 2$ . Then, for every  $g \in \mathcal{G}_0$ :

$$h_\mu(g, T) \geq \text{Cs}(g) \cdot \log M.$$

**Proof.** Consider a shift  $\sigma$  over the alphabet  $\mathcal{A} = \{0, 1, \dots, M - 1\}$  with the corresponding Bernoulli measure generated by  $p_1 = \dots = p_M = \frac{1}{M}$ . It is easy to see that  $h_\mu(\sigma) = \log M$ . From Theorem 4, we conclude that  $\sigma$  is a factor of  $T$ . Therefore, applying formula (10), we obtain:

$$h_\mu(g, T) \geq h_\mu(g, \sigma) \geq h(g, \sigma, \mathcal{P}^{\mathcal{A}}) = \limsup_{n \rightarrow \infty} \frac{1}{n} \varphi(M^{-n}) = \log M \cdot \limsup_{n \rightarrow \infty} \frac{\varphi(M^{-n})}{\varphi_\eta(M^{-n})}.$$

Applying Lemma 8 completes the proof.  $\square$

### 2.2.2. Main Theorem

Our goal in this section is the following result:

**Theorem 5.** Let  $T$  be an ergodic automorphism of Lebesgue space  $(X, \Sigma, \mu)$ , and  $g \in \mathcal{G}_0$  be such that  $\text{Cs}(g) \in (0, \infty)$ . Then:

$$h_\mu(g, T) = \begin{cases} \text{Cs}(g) \cdot h_\mu(T), & \text{if } h_\mu(T) < \infty, \\ \infty, & \text{otherwise.} \end{cases}$$

If  $g \in \mathcal{G}_0^0$ , then  $h_\mu(g, T) = 0$ . If  $g \in \mathcal{G}_0$  is such that  $\text{Cs}(g) = \infty$  and  $T$  has a positive measure-theoretic entropy, then  $h_\mu(g, T) = \infty$ .

Moreover, for  $g \in \mathcal{G}_0^\infty$  from Theorem 3, we have:

**Corollary 5.** Let  $g \in \mathcal{G}_0^\infty$ . If  $(X, T)$  is aperiodic and surjective, then  $h_\mu(g, T) = \infty$ .

To prove Theorem 5, we need the first few preliminary lemmas.

**Lemma 9.** If  $T$  is an automorphism of the Lebesgue space  $(X, \Sigma, \mu)$ , then for every  $g \in \mathcal{G}_0$ :

$$h_\mu(g, T^m) \leq m h_\mu(g, T).$$

**Proof.** Let  $\mathcal{P} \in \mathfrak{B}$ ,  $m \in \mathbb{N}$ . We have:

$$\begin{aligned} h(g, T, \mathcal{P}) &= \limsup_{k \rightarrow \infty} \frac{1}{k} H(g, \mathcal{P} \vee T^{-m} \mathcal{P} \vee \dots \vee T^{-m(k-1)} \mathcal{P}) \\ &= \lim_{n \rightarrow \infty} \sup_{k \geq n} \frac{1}{k} H(g, \mathcal{P} \vee T^{-m} \mathcal{P} \vee \dots \vee T^{-m(k-1)} \mathcal{P}). \end{aligned}$$

Fix  $k \in \mathbb{N}$ . Then,  $\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}$  is a refinement of  $\mathcal{P} \vee T^{-m}\mathcal{P} \vee \dots \vee T^{-m(k-1)}\mathcal{P}$  for  $n = km, \dots, k(m+1) - 1$ .

Therefore:

$$\begin{aligned} \frac{1}{k}H(g, \mathcal{P} \vee T^{-m}\mathcal{P} \vee \dots \vee T^{-m(k-1)}\mathcal{P}) &\leq \frac{1}{k}H(g, \mathcal{P}_{n-1}) = \frac{n}{k} \frac{1}{n}H(g, \mathcal{P}_{n-1}) \\ &\leq \frac{km + m - 1}{k} \frac{1}{n}H(g, \mathcal{P}_{n-1}) \\ &\leq m \left(1 + \frac{1}{k}\right) \frac{1}{n}H(g, \mathcal{P}_{n-1}) \end{aligned} \tag{12}$$

for  $n = km, \dots, k(m + 1) - 1$ . Let us introduce the following notation:

$$c_k := \frac{1}{k}H(g, \mathcal{P} \vee T^{-m}\mathcal{P} \vee \dots \vee T^{-m(k-1)}\mathcal{P}), \quad a_n := \frac{1}{n}H(g, \mathcal{P}_{n-1}).$$

Then, we can rewrite Equation (12) in the form:

$$c_k \leq m \left(1 + \frac{1}{k}\right) a_n \tag{13}$$

for  $n = km, \dots, km + m - 1$ . Taking the supremum in Equation (13), we obtain:

$$\sup_{l \geq k} c_l \leq m \left(1 + \frac{1}{k}\right) \sup_{n=lm, \dots, l(m+1)-1} a_n \leq m \left(1 + \frac{1}{k}\right) \sup_{n \geq km} a_n.$$

Therefore:

$$\limsup_{k \rightarrow \infty} c_k \leq m \limsup_{n \rightarrow \infty} a_n,$$

and this is equivalent to the statement:

$$h(g, T^m, \mathcal{P}) \leq mh(g, T, \mathcal{P}).$$

Taking the supremum over all finite partitions, we obtain the assertion.  $\square$

Next, the lemma will be just a weaker version of Theorem 5.

**Lemma 10.** *If an automorphism  $T^m$  of a Lebesgue space  $(X, \Sigma, \mu)$  is ergodic for every  $m \in \mathbb{N}$ , then for every  $g \in \mathcal{G}_0$ , such that  $Cs(g) < \infty$  holds:*

$$h_\mu(g, T) = Cs(g) \cdot h_\mu(T).$$

*If  $g \in \mathcal{G}_0^0$ , then  $h_\mu(g, T) = 0$ . If  $g \in \mathcal{G}_0$  is such that  $Cs(g) = \infty$  and  $T$  has a positive Kolmogorov-Sinai entropy, then  $h_\mu(g, T) = \infty$ .*

**Proof.** The case of  $g \in \mathcal{G}_0^0$  follows from Corollary 2. Suppose that there exists such  $g \in \mathcal{G}_0 \setminus \mathcal{G}_0^0$ , which fulfills the assumptions of the lemma and for which we have:

$$Cs(g) \cdot h_\mu(T) - h_\mu(g, T) > 0.$$

Then, applying Lemma 9 to the transformation  $T^m$  and using equality  $h_\mu(T^m) = mh_\mu(T)$  (see [2], Theorem 4.3.16), we obtain:

$$Cs(g)h_\mu(T^m) - h_\mu(g, T^m) \geq m (Cs(g)h_\mu(T) - h_\mu(g, T)) \rightarrow \infty \text{ as } m \rightarrow \infty.$$

Therefore, for sufficiently large  $m$ , there exists an integer  $M$  for which:

$$h_\mu(g, T^m) \leq mh_\mu(g, T) < Cs(g) \log M \leq m Cs(g)h_\mu(T) = Cs(g)h_\mu(T^m). \tag{14}$$

Proposition 2 applied to the transformation  $T^m$  guarantees that for every  $g \in \mathcal{G}_0$  with positive (finite)  $Cs(g)$ , we have:

$$h_\mu(g, T^m) \geq Cs(g) \log M. \tag{15}$$

Comparing Equations (14) and (15), we obtain the contradiction, which implies that

$$h_\mu(g, T) = Cs(g)h_\mu(T).$$

If  $Cs(g) = \infty$  and  $h_\mu(T) > 0$ , then there exists such an integer  $m > 0$  that:

$$h_\mu(T^m) = mh_\mu(T) > \log M$$

and by Proposition 2 and Lemma 9:

$$h_\mu(g, T) = h_\mu(g, T^m) = \infty$$

which completes the proof.  $\square$

**Proof of Theorem 5.** If  $h_\mu(T) = 0$ , the statement is true, because for any  $\mathcal{P} \in \mathfrak{B}$ , we have:

$$0 \leq h(g, \mathcal{P}) \leq Cs(g)h(\mathcal{P}) = 0.$$

Suppose that  $0 < h_\mu(T) < \infty$ . Automorphism  $T$  is ergodic. Thus, it has a factor, which is a Bernoulli automorphism  $T'$  with  $h_\mu(T) = h_\mu(T')$ . Every Bernoulli automorphism is mixing, so  $T^m$  is ergodic for each  $m$ . Applying Lemma 10, we obtain:

$$h_\mu(g, T') = Cs(g)h_\mu(T') = Cs(g)h_\mu(T).$$

Since  $T'$  is a factor of  $T$ , Proposition 1 implies that:

$$Cs(g)h_\mu(T) = Cs(g)h_\mu(T') = h_\mu(g, T') \leq h_\mu(g, T) \leq Cs(g)h_\mu(T)$$

which completes the proof of the case of finite  $h_\mu(T)$ . If  $h_\mu(T) = \infty$ , then Proposition 2 implies that:

$$h_\mu(g, T) \geq Cs(g) \log M$$

for every  $M > 0$ .  $\square$

### 2.2.3. Generator Theorem Counterpart

In the case of  $g \in \mathcal{G}_0^\infty$ , there is no counterpart of a Kolmogorov-Sinai generator theorem, which states that the measure-theoretic entropy of the transformation  $T$  is realized on every generator of the  $\sigma$ -algebra  $\Sigma$ . Let us consider Sturm shifts, shifts that model translations of the circle  $\mathbb{T} = [0, 1)$ . Let  $\beta \in [0, 1)$  and consider the translation  $\phi_\beta: [0, 1) \mapsto [0, 1)$  defined by  $\phi_\beta(x) = x + \beta \pmod{1}$ . Let  $\mathcal{P}$  denote the partition of  $[0, 1)$  given by  $\mathcal{P} = \{[0, \beta), [\beta, 1)\}$ . Then, we associate a binary sequence to each  $t \in [0, 1)$  according to its itinerary relative to  $\mathcal{P}$ ; that is, we associate to  $t \in [0, 1)$  the bi-infinite sequence  $x$  defined by  $x_i = 0$  if  $\phi_\beta^i(t) \in [0, \beta)$  and  $x_i = 1$  if  $\phi_\beta^i(t) \in [\beta, 1)$ . The set of such sequences is not necessary closed, but it is shift-invariant and so its closure is a shift space called the Sturmian shift. If  $\beta$  is irrational, then Sturmian shift is minimal, *i.e.*, there is no proper subshift. Moreover, for a minimal Sturmian shift, the number of  $n$ -blocks that occur in an infinite shift space is exactly  $n + 1$ . Therefore for zero-coordinate partition  $\mathcal{P}^A$ , which is a finite generator of  $\sigma$ -algebra  $\Sigma$ , and for any function  $g \in \mathcal{G}_0$ , we have:

$$H(g, \mathcal{P}_n^A) = \sum_{A \in \mathcal{P}_n^A} g(\mu_S(A)) \leq \varphi \left( \frac{1}{n+1} \right)$$

where  $\mu_S$  is the unique invariant measure for Sturm shift. Thus,

$$h(g, \mathcal{P}^A) \leq \limsup_{n \rightarrow \infty} \frac{n+1}{n} g \left( \frac{1}{n+1} \right) = 0.$$

On the other hand, since it is strictly ergodic (thus aperiodic), Theorem 3 implies that for any  $g \in \mathcal{G}_0^\infty$ :

$$h_\mu(g, T) = \infty.$$

Therefore, we have a finite generator, for which the supremum is not attained.

### 3. Discussion

In this note, we discussed the generalization of the dynamical and Kolmogorov-Sinai entropy based on the idea of considering in the definition of the dynamical entropy a concave function vanishing at the origin instead of the Shannon entropy function  $\eta$ . The connections between dynamical entropies and  $g$ -entropies show that the crucial property of  $\eta$  for applications of the Shannon entropy in the dynamical systems is the behavior of  $\eta$  in the neighborhood of zero. Additionally, the main result of the paper obtained for the generalization of the KS entropy states that, usually, there is a linear dependence between the obtained invariant and the KS entropy. It also implies (due to the fact that it is a continuous function of the entropy) that the measure-theoretic  $g$ -entropy is finitely observable (see, e.g., [23]). Moreover, considering functions that behave in the neighborhood of zero differently than the Shannon function usually trivializes the theory. On the other hand, we showed that if the limit of  $g/\eta$  at zero is infinite, then for every positive number  $\gamma$ , there exists a partition for which the  $g$ -entropy will be greater than

or equal to  $\gamma$ . Thus, the measure-theoretic  $g$ -entropy will be infinite. The example from Section 2.2.3, based on this result, implies that there is no counterpart of the generator theorem (e.g., Tsallis entropies for  $\alpha < 1$ , which we obtain considering  $g(x) = \frac{x-x^\alpha}{\alpha-1}$  with  $\alpha \in (0, 1)$ , fit into this scheme). However, the concept of  $g$ -entropies is still of use, e.g., considering the rate of convergence of partial  $g$ -entropies may give additional information about the system [30]. The most promising direction in this context seems to be considering functions for which  $C(g) = 0$  and  $g'(0) = \infty$ .

## Conflicts of Interest

The author declares no conflict of interest.

## References

1. Downarowicz, T. *Entropy in Dynamical Systems*; Cambridge University Press: New York, NY, USA, 2011.
2. Katok, A.; Hasselblatt, B. *Introduction to the Modern Theory of Dynamical Systems*; Cambridge University Press: New York, NY, USA, 1997.
3. Misiurewicz, M. A short proof of the variational principle for  $\mathbb{Z}_+^n$  action on a compact space. *Astérisque* **1976**, *40*, 147–157.
4. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656.
5. Rényi, A. On measures of entropy and information. In Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, Statistical Laboratory of the University of California, Berkeley, CA, USA, 20 June–30 July 1960; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 547–561.
6. Arimoto, S. Information-theoretical considerations on estimation problems. *Inf. Control* **1971**, *19*, 181–194.
7. Wu, Y.; Verdú, S. Rényi Information Dimension: Fundamental Limits of Almost Lossless Analog Compression. *IEEE Trans. Inf. Theory* **2010**, *56*, 3721–3748.
8. Csiszár, I. Axiomatic characterization of information measures. *Entropy* **2008**, *10*, 261–273.
9. De Paly, T. On entropy-like invariants for dynamical systems. *Z. Anal. Anwend.* **1982**, *1*, 69–79.
10. De Paly, T. On a class of generalized K-entropies and Bernoulli shifts. *Z. Anal. Anwend.* **1982**, *1*, 87–96.
11. Grassberger, P.; Procaccia, I. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev. A* **1983**, *28*, 2591–2593.
12. Takens, F.; Verbitski, E. Generalized entropies: Rényi and correlation integral approach. *Nonlinearity* **1998**, *11*, 771–782.
13. Takens, F.; Verbitski, E. Rényi entropies of aperiodic dynamical systems. *Isr. J. Math.* **2002**, *127*, 279–302.

14. Liu, Q.; Cao, K.-F.; Peng, S.-L. A generalized Kolmogorov–Sinai-like entropy under Markov shifts in symbolic dynamics. *Physica A* **2009**, *388*, 4333–4344.
15. Mesón, A.M.; Vericat, F. Invariant of dynamical systems: A generalized entropy. *J. Math. Phys.* **1996**, *37*, 4480–4483.
16. Mesón, A.M.; Vericat, F. On the Kolmogorov-like generalization of Tsallis entropy, correlation entropies and multifractal analysis. *J. Math. Phys.* **2002**, *43*, 904–918.
17. Havrda, J.; Charvát, F. Quantification method of classification processes. Concept of structural  $\alpha$ -entropy. *Kybernetika* **1967**, *3*, 30–35.
18. Abe, S. Tsallis entropy: How unique? *Contin. Mech. Thermodyn.* **2004**, *16*, 237–244.
19. Furuichi, S. Information theoretical properties of Tsallis entropies. *J. Math. Phys.* **2006**, *47*, 023302.
20. Tsallis, C. Possible generalization of Boltzmann–Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487.
21. Tsallis, C. Entropic nonextensivity: A possible measure of complexity. *Chaos Solitons Fractals* **2002**, *13*, 371–391.
22. Tsallis, C.; Plastino, A.R.; Zheng, W.-M. Power-law sensitivity to initial conditions—New entropic representation. *Chaos Solitons Fractals* **1997**, *8*, 885–891.
23. Ornstein, D.S.; Weiss, B. Entropy is the only finitely observable invariant. *J. Mod. Dyn.* **2007**, *1*, 93–105.
24. Weiss, B. (The Hebrew University of Jerusalem), **2013**, personal communication.
25. Weiss, B. *Single Orbit Dynamics*; American Mathematical Society: Providence, RI, USA, 2000.
26. Blume, F. The Rate of Entropy Convergence. Ph.D. Thesis, University of North Carolina, Chapel Hill, NC, USA, 1995.
27. Blume, F. Possible rates of entropy convergence. *Ergod. Theory Dyn. Syst.* **1997**, *17*, 45–70.
28. Galatolo, S. Global and local complexity in weakly chaotic dynamical systems. *Discret. Contin. Dyn. Syst.* **2003**, *9*, 1607–1624.
29. Ferenczi, S.; Park, K.K. Entropy dimensions and a class of constructive examples. *Discret. Contin. Dyn. Syst.* **2007**, *17*, 133–141.
30. Falniowski, F. Possible  $g$ -entropy convergence rates. **2013**, arXiv:1309.6246.
31. Rosenbaum, R.A. Sub-additive functions. *Duke Math. J.* **1950**, *17*, 227–247.
32. Heinemann, S.-M.; Schmitt, O. Rokhlin’s Lemma for non-invertible maps. *Dyn. Syst. Appl.* **2001**, *10*, 201–214.
33. Sinai, Y.G. Weak isomorphism of transformation with an invariant measure. *Sov. Math.* **1962**, *3*, 1725–1729.