

# RÉNYI ENTROPIES OF APERIODIC DYNAMICAL SYSTEMS

BY

FLORIS TAKENS

*Department of Mathematics, University of Groningen  
P.O.Box 800, 9700 AV, Groningen, The Netherlands  
e-mail: f.takens@math.rug.nl*

AND

EVGENY VERBITSKIY

*Eurandom, P.O.Box 513, 5600 MB, Eindhoven, The Netherlands  
e-mail: e.verbitskiy@tue.nl*

## ABSTRACT

In this paper we continue the study of Rényi entropies of measure-preserving transformations started in [22]. We have established there that for ergodic transformations with positive entropy, the Rényi entropies of order  $q$ ,  $q \in \mathbb{R}$ , are equal to either plus infinity ( $q < 1$ ), or to the measure-theoretic (Kolmogorov–Sinai) entropy ( $q \geq 1$ ). The answer for non-ergodic transformations is different: the Rényi entropies of order  $q > 1$  are equal to the essential infimum of the measure-theoretic entropies of measures forming the decomposition into ergodic components. Thus, it is possible that the Rényi entropies of order  $q > 1$  are strictly smaller than the measure-theoretic entropy, which is the average value of entropies of ergodic components.

This result is a bit surprising: the Rényi entropies are metric invariants, which are sensitive to ergodicity.

The proof of the described result is based on the construction of partitions with independent iterates. However, these partitions are obtained in different ways depending on  $q$ : for  $q > 1$  we use a version of the well-known Sinai theorem on Bernoulli factors for the non-ergodic transformations; for  $q < 1$  we use the notion of collections of independent sets in Rokhlin–Halmos towers and their properties.

---

Received December 8, 1999 and in revised form January 4, 2001

## 1. Introduction

Alfred Rényi introduced the generalization of the Shannon information (entropy) in the beginning of sixties. His approach was based on an axiomatic definition of information, and consisted of including the standard entropy function

$$H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log p_i$$

into a one-parameter family of generalized entropy functions

$$H_q(p_1, \dots, p_n) = - \frac{1}{q-1} \log \left( \sum_{i=1}^n p_i^q \right), \quad q \neq 1.$$

For a fixed probability distribution  $(p_1, \dots, p_n)$  the standard entropy is recovered from the generalized entropies as follows,

$$H(p_1, \dots, p_n) = \lim_{q \rightarrow 1} H_q(p_1, \dots, p_n).$$

Since then the Rényi entropies have been successfully used in information theory and statistics, and more recently in thermodynamics and quantum mechanics. In dynamical systems, Hentschel and Procaccia [8] suggested a one-parameter family of generalized dimensions based on Rényi's approach. These dimensions proved to be extremely useful in problems of multifractal analysis and characterization of chaotic attractors; see, e.g., [13].

Some attempts [7], [6] were made to introduce the generalized entropies of dynamical system using Rényi's approach. The idea was to produce a sufficiently rich family of invariants of a dynamical system, which will take into account the non-uniform behavior of invariant measures. However, the proposed way of generalizing the Kolmogorov–Sinai entropy using  $H_q$  instead of  $H_1$  turned out to be non-productive. In [22] we have established the following fact.

**THEOREM 1.1:** *For an ergodic dynamical system  $(X, \mathfrak{B}, \mu, T)$  with positive measure-theoretic entropy  $h(T, \mu) > 0$ , the Rényi entropies are given by the following formula:*

$$h(T, q, \mu) = \begin{cases} +\infty, & q < 1, \\ h(T, \mu), & q \geq 1. \end{cases}$$

Also in [22] we suggested another family of generalized entropies, which recovers the results reported in the physics literature [1].

The proof of Theorem 1.1 relies heavily on the Sinai theorem on Bernoulli factors [19], for which the assumptions of ergodicity and positiveness of the measure-theoretic entropy are crucial.

In this paper we prove a result, similar to Theorem 1.1, but without the above assumptions. We consider aperiodic measure-preserving automorphisms, i.e., transformations  $T$  defined on Lebesgue space  $(X, \mathfrak{B}, \mu)$  such that

$$\mu(\{x: T^n(x) = x \text{ for some } n\}) = 0.$$

The result for such systems is different from the ergodic case.

**THEOREM 1.2:** *Suppose  $T$  is an aperiodic measure-preserving automorphism of the Lebesgue space  $(X, \mathfrak{B}, \mu)$ . Let  $\mu = \int \mu_t dm(t)$  be the decomposition of  $\mu$  into ergodic components, and let*

$$h_*(T, \mu) = \text{m-essinf}\{h(T, \mu_t)\} := \sup\left\{c: m\{t: h(T, \mu_t) < c\} = 0\right\}.$$

*Then the Rényi entropies are as follows:*

$$h(T, q, \mu) = \begin{cases} +\infty, & q < 1, \\ h(T, \mu) = \int h(T, \mu_t) dm(t), & q = 1, \\ h_*(T, \mu), & q > 1. \end{cases}$$

This result is a bit surprising because of the following: an entropy-based invariant can detect ergodicity. However, we are not aware of any interesting example where this observation could be useful. The first candidates which come to mind are the non-ergodic Markov shifts, i.e., the shifts for which the transition probability matrix  $P$  is not irreducible. It is possible in this case (provided  $h(T, \mu) > h_*(T, \mu)$ , of course) to show the Rényi entropies of order  $q > 1$  are strictly smaller than the measure-theoretic entropy, and thus the system is not ergodic. However, this proof is much more involved than the standard one and follows the same idea.

The paper is organized as follows: in the next section we give a formal definition of the Rényi entropies and establish the basic properties; in section 3 we recall facts about the decomposition into ergodic components. We discuss a non-ergodic version of Sinai's theorem on Bernoulli factors, and use it for the computation of the Rényi entropies of order  $q > 1$  in section 4. In section 5 we develop a notion of independent partitions in Rokhlin–Halmos towers and subsequently prove the statement for  $q < 1$ . Finally, in the last section, we pose some open questions about the possible connection between the Rényi entropies and the recently introduced entropy convergence rates.

## 2. Rényi entropies of measure-preserving transformations

The definition of the Rényi entropy of order  $q$  of a measure-preserving transformation goes along the lines of the standard definition of the measure-theoretic (Kolmogorov–Sinai) entropy and consists of 3 steps: the definition of the Rényi entropy of a finite partition, Rényi entropy of an automorphism with respect to a partition, and, finally, after taking the supremum over all finite partitions, the Rényi entropy of an automorphism, which is a metric invariant.

For any  $q \in \mathbb{R}$  the entropy of order  $q$  of the partition  $\xi = \{\Delta_i\}_{i=1}^n$  is the number

$$(2.1) \quad H_\mu(q, \xi) = \begin{cases} -\frac{1}{q-1} \log \left( \sum_{i=1}^n \mu(\Delta_i)^q \right), & \text{for } q \neq 1, \\ -\sum_{i=1}^n \mu(\Delta_i) \log \mu(\Delta_i), & \text{for } q = 1, \end{cases}$$

with the standard convention  $0^q = 0$  for all  $q \in \mathbb{R}$  and  $0 \log 0 = 0$ .

It is easy to check the following monotonicity property:

$$H_\mu(q_1, \xi) \leq H_\mu(q_2, \xi) \quad \text{for any } \xi \text{ and } q_1 \geq q_2.$$

The Rényi entropy of order  $q$  with respect to a partition  $\xi$  is defined as

$$(2.2) \quad h(T, \mu, q, \xi) = \liminf_{n \rightarrow \infty} \frac{1}{n} H_\mu(q, \xi^{(n)}),$$

where  $\xi^{(n)} = \xi \vee T^{-1}\xi \vee \dots \vee T^{-n+1}\xi$  is the partition into sets  $\bigcap_{k=0}^{n-1} T^{-k}\Delta_{i_k}$  with  $\Delta_{i_k} \in \xi$ .

*Remark:* For  $q = 1$  it is known (see for example [4]) that the limit in (2.2) exists. The proof of this fact is based on a so-called **subadditivity property** of the Shannon entropy  $H(1, \xi)$ :

$$H_\mu(1, \xi \vee \eta) \leq H_\mu(1, \xi) + H_\mu(1, \eta)$$

for all partitions  $\xi, \eta$ . As was shown by Rényi in [14], the latter is not the case for any  $q \neq 1$ . This creates some additional problems in the treatment of the Rényi entropies. Nevertheless, if  $\xi$  and  $\eta$  are independent partitions then

$$H_\mu(q, \xi \vee \eta) = H_\mu(q, \xi) + H_\mu(q, \eta)$$

for all  $q \in \mathbb{R}$ . We will often exploit this fact.

Finally, we define the Rényi entropy of an automorphism  $T$  of order  $q$  as the number

$$(2.3) \quad h(T, \mu, q) = \sup_{\xi} h(T, \mu, q, \xi),$$

where the supremum is taken over all finite partitions  $\xi$  of  $X$ .

PROPOSITION 2.1: *The Rényi entropies have the following properties:*

- (1)  $h(T, \mu, q) \geq 0$  for all  $q$ ;
- (2)  $h(T, \mu, q_1) \geq h(T, \mu, q_2)$  for  $q_1 \leq q_2$ ;
- (3)  $h(T, \mu, 1) = h(T, \mu)$ , where  $h(T, \mu)$  is the measure-theoretic (or Kolmogorov–Sinai) entropy.
- (4)  $h(T^n, \mu, q) = nh(T, \mu, q)$  for any  $q \in \mathbb{R}$  and every  $n \geq 0$ .

Properties 1–3 follow easily from the definition of  $h(T, \mu, q)$ , and (4) has been established in [22].

### 3. Decomposition into ergodic components

Let  $(X, \mathfrak{B}, \mu)$  be a Lebesgue space [4]. For a measurable partition  $\xi = \{C_t\}_{t \in \Lambda}$ , where  $\Lambda$  can be finite, countable or uncountable, we identify  $\Lambda$  and the quotient (or factor)  $X/\xi$  — the space whose points are the elements of  $\xi$ . The set  $\Lambda$  is a Lebesgue space as well: the set  $E \subset \Lambda$  is measurable if the set  $\bigcup_{t \in E} C_t$  is a measurable subset of  $X$ , and we obtain a measure  $m$  on  $\Lambda$  by letting  $m(E) = \mu(\bigcup_{t \in E} C_t)$ . A system of measures  $\{\mu_t\}$ ,  $t \in \Lambda$ , is called a **canonical system of conditional measures belonging to the partition**  $\xi = \{C_t\}_{t \in \Lambda}$ , if

- (1)  $\mu_t$  is defined on some  $\sigma$ -algebra  $\mathfrak{B}_t$  of subsets of  $C_t$ , such that  $(C_t, \mathfrak{B}_t, \mu_t)$  is a Lebesgue space,
- (2) for any  $A \in \mathfrak{B}$  the set  $A \cap C_t$  belongs to  $\mathfrak{B}_t$  for  $m$ -almost all  $t$ ; the function  $\mu_t(A \cap C_t)$  is a measurable function of  $t$  and

$$\mu(A) = \int \mu_t(A \cap C_t) dm(t).$$

Suppose  $T: X \rightarrow X$  is a measure-preserving automorphism. Then  $(X, \mathfrak{B}, \mu)$  can be decomposed into ergodic components of  $T$ . By this we mean the following: there exists a  $T$ -invariant measurable partition  $\xi = \{C_t\}$  and a canonical system of conditional measures  $\{\mu_t\}$  such that for almost all  $t$

$$(C_t, \mathfrak{B}_t, \mu_t, T|_{C_t}) \text{ is ergodic.}$$

Suppose  $\xi = \{C_t\}$  is the decomposition into ergodic components of  $(X, \mathfrak{B}, \mu, T)$ ; then

$$h(T, \mu) = \int h(T, \mu_t) dm.$$

Consider the essential infimum and the essential supremum of measure-theoretic

entropies of the measures  $\mu_t$  from the decomposition into ergodic components:

$$\begin{aligned} h_*(T, \mu) &= \text{m-essinf}\{h(T, \mu_t) \mid t \in \Lambda = X/\xi\} \\ &:= \sup\{c : m(\{t : h(T, \mu_t) < c\}) = 0\}, \\ h^*(T, \mu) &= \text{m-esssup}\{h(T, \mu_t) \mid t \in \Lambda = X/\xi\} \\ &:= \inf\{c : m(\{t : h(T, \mu_t) > c\}) = 0\}. \end{aligned}$$

The quantity  $h^*(T, \mu)$ , sometimes called the **entropy rate**, has been previously studied in the literature [9, 21, 23] in relation to the existence of finite generators (generating partitions) for non-ergodic systems. A well-known theorem of Krieger [11] states that an ergodic dynamical system with a finite measure-theoretic entropy  $h(T, \mu)$  admits a finite generator  $\xi$  with  $\text{card}(\xi) \leq \exp(h(T, \mu)) + 1$ . It turns out that for non-ergodic aperiodic dynamical systems a similar result is true, provided  $h^*(T, \mu) < \infty$ : a finite generator  $\xi$  exists whose cardinality does not exceed  $\exp(h^*(T, \mu)) + 1$ .

Denote by  $\Pi_m = \{P = (P_1, \dots, P_m)\}$  the set of all ordered partitions of  $X$  into  $m$  sets. For any measure  $\mu$  on  $(X, \mathfrak{B})$  define the **partition (pseudo-)metric**  $\rho_\mu$  on  $\Pi_m$  as follows:

$$\rho_\mu(P, Q) = \sum_{k=1}^m \mu(P_k \Delta Q_k), \quad P, Q \in \Pi_m.$$

If  $\rho_\mu(P, Q) = 0$  then  $P$  and  $Q$  agree except on a set of measure 0 and, of course, in this case we say that  $P = Q$ . The space  $(\Pi_m, \rho_\mu)$  is a complete metric space.

For an at most countable ordered partition  $P$  of  $(X, \mathfrak{B}, \mu)$  the distribution vector of  $P$  is given by

$$d(P, \mu) = (\mu(P_1), \mu(P_2), \dots).$$

Suppose  $P$  and  $\bar{P}$  are partitions into  $m$  sets of  $(X, \mathfrak{B}, \mu)$ ,  $(Y, \mathcal{F}, \nu)$  respectively; then the **distribution distance** is

$$|d(P, \mu) - d(\bar{P}, \nu)| := \sum_{k=1}^m |\mu(P_k) - \nu(\bar{P}_k)|.$$

Suppose we have a set  $\{\mu_t\}_{t \in \Lambda}$  of measures on  $(X, \mathfrak{B})$ . For every  $t \in \Lambda$  consider the metric  $\rho_{\mu_t}$  on  $\Pi_m$ . The following fact will be used later: there exists a countable set  $\tilde{\Pi}_m \subset \Pi_m$  which is  $\rho_{\mu_t}$ -dense in  $\Pi$  for almost every  $t \in \Lambda$ .

The existence of such  $\tilde{\Pi}_m$  follows from the fundamental properties of the Lebesgue spaces. By definition, a Lebesgue space  $(X, \mathfrak{B}, \mu)$  admits a countable

basis  $\Gamma = \{B_\alpha\}$ . This in particular means that for any measurable set  $A \in \mathfrak{B}$  there exists a set  $C$  from a minimal  $\sigma$ -algebra generated by  $\Gamma$  such that

$$(3.1) \quad C \subset A \quad \text{and} \quad \mu(A \setminus C) = 0.$$

Denote by  $\mathfrak{A}$  the countable algebra generated by  $\Gamma$ , and let

$$\tilde{\Pi}_m = \left\{ P = (P_1, \dots, P_m) : P_i \in \mathfrak{A} \right\}.$$

Hence  $\tilde{\Pi}_m$  is an at most countable collection of ordered partitions into  $m$  sets, where elements of these partitions are taken from  $\mathfrak{A}$ . From (3.1) we conclude that  $\tilde{\Pi}_m$  is  $\rho_\mu$  dense in  $\Pi_m$ . Moreover, for almost every  $t \in \Lambda$ ,  $\tilde{\Pi}_m$  is  $\rho_{\mu_t}$ -dense in  $\Pi_m$  as well. This is a consequence of the following fact ([15], see also [16]): for almost every  $t \in \Lambda$ , the countable collection of sets  $\Gamma_t = \Gamma \cap C_t$  is a basis in the Lebesgue space  $(C_t, \mathfrak{B}_t, \mu_t)$ .

#### 4. Rényi entropies of order $q > 1$

In this section we are going to prove that  $h(T, \mu, q) = h_*(T, \mu)$  for every  $q > 1$ . We start by showing that  $h(T, \mu, q) \leq h_*(T, \mu)$ .

**4.1. ESTIMATE FROM ABOVE.** Suppose that we have two invariant measures  $\mu_1$  and  $\mu_2$  for an automorphism  $T$ . We do not assume these measures to be ergodic. Without loss of generality we can assume that

$$(4.1) \quad h(T, \mu_1) \leq h(T, \mu_2).$$

Consider now another invariant measure  $\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$  with  $\alpha \in (0, 1)$ . The measure-theoretic entropy of  $\mu$  is given by (see [5])

$$h(T, \mu) = \alpha h(T, \mu_1) + (1 - \alpha)h(T, \mu_2).$$

Note that due to (4.1),  $h(T, \mu) \geq h(T, \mu_1)$ . Let  $\xi$  be some finite partition. For any  $C \in \xi$  one has

$$\mu(C) = \alpha\mu_1(C) + (1 - \alpha)\mu_2(C),$$

and, therefore,  $\mu(C)^q \geq \alpha^q\mu_1(C)^q$  for  $q > 1$ . Hence, for  $q > 1$

$$\begin{aligned} H_\mu(q, \xi) &= -\frac{1}{q-1} \log \left( \sum_{C \in \xi} \mu(C)^q \right) \leq -\frac{q}{q-1} \log \alpha - \frac{1}{q-1} \log \left( \sum_{C \in \xi} \mu_1(C)^q \right) \\ &= -\frac{q}{q-1} \log \alpha + H_{\mu_1}(q, \xi). \end{aligned}$$

From the above one easily concludes that

$$h(T, \mu, q, \xi) = \liminf_{n \rightarrow \infty} \frac{1}{n} H_\mu(q, \xi^{(n)}) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} H_{\mu_1}(q, \xi^{(n)}) = h(T, \mu_1, q, \xi).$$

On the other hand, due to the monotonicity of the Rényi entropies with respect to  $q$ , for  $q > 1$  we have

$$h(T, \mu_1, q, \xi) \leq h(T, \mu_1, q) \leq h(T, \mu_1, 1) = h(T, \mu_1).$$

Combining the two last inequalities we finally obtain that, for any  $q > 1$ ,

$$h(T, \mu, q) \leq h(T, \mu_1).$$

Thus we see that the Rényi entropy of a linear combination of two measures does not exceed a minimum of the measure-theoretic entropies of these two measures. It is evident that the above argument goes through in the case of a finite or countable decomposition:  $\mu = \sum_k \alpha_k \mu_k$ , where  $\alpha_k \geq 0$  and  $\sum_k \alpha_k = 1$ .

Moreover, the above argument can be equally easily generalized to the case of generally uncountable decomposition of an invariant measure  $\mu$  into ergodic components  $\{\mu_t\}$ . This is done in the following lemma.

**LEMMA 4.1:** *For a measure preserving system  $(X, \mathfrak{B}, \mu, T)$  one has*

$$(4.2) \quad h(T, \mu, q) \leq h_*(T, \mu)$$

for every  $q > 1$ .

*Proof:* Consider an ergodic decomposition of  $(X, \mathfrak{B}, \mu, T)$  as in section 3. By the definition of  $h_*(T, \mu)$ , for every  $\varepsilon > 0$  the set  $E_1 = \{t: h(T, \mu_t) < h_*(T, \mu) + \varepsilon\}$  has a positive  $m$ -measure. Suppose there exists  $\varepsilon_0 > 0$  such that for any  $\varepsilon \in (0, \varepsilon_0)$  one has  $m(E_1) < 1$ .

If such  $\varepsilon_0 > 0$  does not exist, then

$$h(T, \mu_t) = h_*(T, \mu) \quad \text{for } m\text{-a.a. } t.$$

As a result we immediately conclude that  $h(T, \mu) = h_*(T, \mu)$ , and using the fact that  $h(T, \mu) \geq h(T, \mu, q)$  for any  $q > 1$  we obtain our claim (4.2).

Assume such  $\varepsilon_0 > 0$  exists and choose any  $\varepsilon \in (0, \varepsilon_0)$ . Since  $m(E_1) \in (0, 1)$  we can define

$$\mu_1 = \frac{1}{m(E_1)} \int_{E_1} \mu_t dm(t), \quad \mu_2 = \frac{1}{1 - m(E_1)} \int_{E_1^c} \mu_t dm(t).$$

It is clear that  $\mu_1$  and  $\mu_2$  are invariant probability measures. Moreover,  $h(T, \mu_1) \leq h_*(T, \mu) + \varepsilon$ . Using the above argument for two measures  $\mu_1$  and  $\mu_2$  we conclude that, for any  $q > 1$ ,

$$h(T, \mu, q) \leq h_*(T, \mu) + \varepsilon.$$

Since  $\varepsilon > 0$  can be chosen arbitrary small, we obtain the claim (4.2).  $\blacksquare$

**4.2. BERNoulli FACTORS OF NON-ERGODIC SYSTEMS.** Let us recall a definition of a **Bernoulli automorphism**.

**Definition 4.2:** An automorphism  $T$  of a Lebesgue space  $(X, \mathfrak{B}, \mu)$  is called Bernoulli, if it is measure-theoretically isomorphic to a Bernoulli shift.

If  $T$  is a Bernoulli automorphism, then there exists a partition  $P$  of  $X$  such that

- (1)  $P$  is generating,
- (2)  $\{T^n P\}_{n \in \mathbb{Z}}$  is a sequence of independent partitions.

Such a partition  $P$  is called an independent generator for  $T$ . A well-known theorem by Sinai [19] states that for every ergodic automorphism  $T$  with entropy  $h(T, \mu)$ , and every positive number  $h$  such that  $h \leq h(T, \mu)$ , there exists a Bernoulli factor with entropy  $h$ . A non-ergodic version of the Sinai theorem first appeared in [10].

**THEOREM 4.3:** Suppose  $T$  is an automorphism of a Lebesgue space  $(X, \mathfrak{B}, \mu)$ . Let  $\bar{T}$  be a Bernoulli automorphism of  $(Y, \mathfrak{F}, \nu)$  with a finite independent generator  $\bar{P}$ ,  $\text{card}(\bar{P}) = k$ . Let  $\{C_t, \mu_t\}$  be a decomposition of  $(X, \mathfrak{B}, \mu, T)$  into ergodic components and  $m$  be a corresponding measure on the factor  $X/\{C_t\}$ . Assume that  $h_*(T, \mu) \geq h(\bar{T}, \nu)$ . Then there exists a partition  $Q$ ,  $\text{card}(Q) = k$ , such that

- (i)  $\{T^i Q\}$  is a sequence of independent partitions,
- (ii)  $d(Q, \mu) = d(\bar{P}, \nu)$ .

In fact, using the techniques of [10] one can establish a non-ergodic version of the Ornstein fundamental lemma [12, 18, 20] as well. The strategy of generalizing “ergodic” results to the non-ergodic case consists of the following. Suppose that  $\{C_t, \mu_t\}$  is the decomposition of  $\mu$  into ergodic components, and that for almost every  $t$  there exists a partition  $P_t$  of  $C_t$  into  $m$  elements which satisfies some required property. We recall that there exists a countable family of partitions  $\tilde{\Pi}_m$  which is  $\rho_{\mu_t}$ -dense in the set all partitions into  $m$  elements for almost all  $t$ ; see section 3. Using this family  $\tilde{\Pi}_m$  one can construct a *universal* partition  $P$  such that  $P \cap C_t = P_t$  for almost all  $t$ .

4.3. ESTIMATE FROM BELOW. Now we can prove a lower estimate:  $h(T, \mu, q) \geq h_*(T, \mu)$  for all  $q > 1$ . Before we proceed with this estimate we would like to make a few remarks. Firstly, we compute the Rényi entropy of order  $q$  for a Bernoulli shift. Let  $\Omega = \{1, \dots, k\}^{\mathbb{Z}}$ , and  $\sigma: \Omega \rightarrow \Omega$  be a left shift. Let  $\nu$  be a Bernoulli measure on  $\Omega$  generated by a probability vector  $p = (p_1, \dots, p_k)$ , i.e.,

$$\nu\{\omega = (\omega_i): \omega_m = a_m, \dots, \omega_n = a_n\} = p_{a_m} \cdots p_{a_n},$$

for all  $m \leq n$  and  $a_m, \dots, a_n \in \{1, \dots, k\}$ . Denote by  $\xi$  the partition into the following cylinders:

$$\xi = \{\Delta_1, \dots, \Delta_k\}, \quad \Delta_n = \{\omega = (\omega_i): \omega_0 = n\}.$$

It is not very difficult to see that, for  $q \neq 1$ ,

$$h(\sigma, \nu, q, \xi) = -\frac{1}{q-1} \log \left( \sum_{i=1}^k p_i^q \right).$$

In particular, if  $p_1 = \dots = p_k = 1/k$ , then  $h(\sigma, \nu, q, \xi) = \log k$ . In this case, since  $h(\sigma, \nu) = \log k$ , we immediately conclude that  $h(\sigma, \nu, q) = \log k$  for  $q > 1$ .

We would also need the following statement.

LEMMA 4.4: *Suppose  $\bar{T}: Y \rightarrow Y$  is an automorphism preserving a measure  $\nu$ . Then for any prime  $p \geq 1$  one has*

$$h_*(\bar{T}^p, \nu) = ph_*(\bar{T}, \nu).$$

*Proof:* Assume first that  $(\bar{T}, \nu)$  is ergodic. If  $\bar{T}^p: Y \rightarrow Y$  is ergodic, then there is nothing to prove, since in this case  $h_*(\bar{T}^p, \nu) = h(\bar{T}^p, \nu) = ph(\bar{T}, \nu)$ .

If  $\bar{T}^p$  is not ergodic, then [17, p. 38] there exist disjoint sets  $A_0, \dots, A_{p-1}$  such that  $Y = \bigcup_{i=0}^{p-1} A_i \pmod{0}$ ,  $\bar{T}(A_i) = A_{i+1 \pmod{p}}$ , and  $\bar{T}^p$  is ergodic on  $A_0$  with respect to  $\nu(\cdot | A_0)$ . Therefore,  $h_*(\bar{T}^p, \nu) = ph(\bar{T}, \nu)$ .

Hence, we conclude that if  $(\bar{T}, \nu)$  is ergodic, then  $h_*(\bar{T}^p, \nu) = ph(\bar{T}, \nu)$  for any prime  $p$ ,  $p \geq 1$ .

Assume now that  $(T, \mu)$  is not ergodic and let  $\mu = \int \mu_t dm$  be the decomposition of  $\mu$  into ergodic components. Applying the argument above to each  $(T, \mu_t)$  we conclude that  $h_*(T^p, \mu_t) = ph(T, \mu_t)$ , and therefore  $h_*(T^p, \mu) = ph_*(T, \mu)$ .

■

Now let us proceed with the proof of the inequality:  $h(T, \mu, q) \geq h_*(T, \mu)$  for all  $q > 1$ . Assume the opposite, i.e., there exists  $q > 1$  such that  $h(T, \mu, q) <$

$h_*(T, \mu)$ . Take a sufficiently large prime  $p$  such that there exists an integer  $k$  satisfying

$$(4.3) \quad ph(T, \mu, q) < \log k < ph_*(T, \mu) = h_*(T^p, \mu).$$

Consider a Bernoulli shift, defined as above, with  $p_1 = \dots = p_k = 1/k$ . Then by Theorem 4.3 there exists a Bernoulli factor  $Q$  for  $T^p$  with  $\mu(Q_1) = \dots = \mu(Q_k) = 1/k$ . Thus  $h(T^p, \mu, q, Q) = \log k$ , but this is in contradiction with (4.3), since

$$ph(T, \mu, q) = h(T^p, \mu, q) = \sup_R h(T^p, \mu, q, R).$$

Hence,  $h(T, \mu, q) \geq h_*(T, \mu)$  for  $q > 1$ , and together with (4.2) this gives the equality  $h(T, \mu, q) = h_*(T, \mu)$  for all  $q > 1$ .

## 5. Rényi entropies of order $q < 1$

In this section we will prove the remaining part of Theorem 1.2. The techniques which we are going to use will be different from the previous section. The reason is that we do not want to assume  $h_*(T, \mu) > 0$  (or even  $h(T, \mu) > 0$ ). In the case when  $h_*(T, \mu) > 0$ , we can (with the help of the non-ergodic version of Sinai's theorem on Bernoulli factors obtained in the previous section) proceed as in [22].

Our main goal is to construct partitions with arbitrarily large Rényi entropy of order  $q$ ,  $q < 1$ : for every  $C > 0$  we have to find a partition  $\xi$  such that

$$(5.1) \quad h(T, \mu, q, \xi) = \liminf_{k \rightarrow \infty} \frac{1}{k} H_\mu(q, \xi^{(k)}) \geq C.$$

Since the Rényi entropies are monotonic in  $q$  we can restrict ourselves to  $q \in (0, 1)$ .

First of all, let us make an observation which will allow us to simplify the estimate of the Rényi entropy of a partition from below.

**Definition 5.1:** The Rényi entropy of order  $q$ ,  $q \neq 1$ , of a finite partition  $\eta = \{\Delta_i\}$  restricted to a set  $F$ ,  $\mu(F) > 0$ , is the number

$$H_\mu(q, \eta|F) = -\frac{1}{q-1} \log \left( \sum_{\Delta_i \in \eta} \mu(\Delta_i \cap F)^q \right).$$

It is easy to see that for each  $q \in (0, 1)$  and for any set  $F$ ,  $\mu(F) > 0$ , one has

$$(5.2) \quad H_\mu(q, \eta) \geq H_\mu(q, \eta|F).$$

In the next subsection we will show how this can be used when  $F$  is a base of some Rokhlin–Halmos tower and  $\xi$  is some special partition.

5.1. **ROKHIN–HALMOS TOWERS AND INDEPENDENT COLLECTIONS OF SETS.** We have assumed that  $T$  is an aperiodic automorphism. It is well known that for such automorphisms one can construct Rokhlin–Halmos towers of any height and measure arbitrarily close to 1.

Let  $M \subset X$ ; then  $\tau = \{M, TM, \dots, T^{m-1}M\}$  is called a **Rokhlin–Halmos tower** if

$$T^i M \cap T^j M = \emptyset \quad \text{for } 0 \leq i \neq j \leq m-1.$$

We will use the same letter  $\tau$  for  $\bigcup_{k=0}^{m-1} T^k M$ . The height of the tower  $\tau$  is said to be  $m$  and  $\mu(\tau) = m\mu(M)$  is its measure.

We now give a definition of an independent collection of sets relative to a Rokhlin–Halmos tower. We will associate to such collections certain partitions, which will be analogous to Bernoulli partitions.

*Definition 5.2:* Let  $\tau = \{M, \dots, T^{m-1}M\}$  be a Rokhlin–Halmos tower. We say that a collection  $\mathcal{I} = \{I_1, \dots, I_{N-1}\}$  of subsets of  $\tau$  is independent relative to  $\tau$  if

- (1)  $I_i \cap I_j = \emptyset$  for  $i \neq j$ ;
- (2) denote by  $\xi_{\mathcal{I}}$  the partition of  $X$  into the sets

$$\{I_1, \dots, I_{N-1}, I_N := X \setminus \bigcup_{j=0}^{N-1} I_j\};$$

then

$$\{T^{-k}(I_1 \cap T^k M), \dots, T^{-k}(I_N \cap T^k M)\}_{k=0}^{m-1} = \{T^{-k}(\xi_{\mathcal{I}} \cap T^k M)\}_{k=0}^{m-1}$$

is a collection of independent partitions of  $M$ .

For convenience we will always assume that

$$(3) \quad \mu(I_j \cap T^k M) = \frac{\mu(M)}{N} \quad \text{for } j = 1, \dots, N \text{ and } k = 0, \dots, m-1.$$

Collections of independent sets exist in every tower. This follows from the following two observations. Firstly, since  $T$  is assumed to be aperiodic, the invariant measure  $\mu$  has no atoms. Secondly, for any Lebesgue space  $(X, \mathfrak{B}, \mu)$ , where  $\mu$  has no atoms, for every measurable set  $A$  and each  $\alpha \in [0, \mu(A)]$  one can find a set  $B \subseteq A$  with  $\mu(B) = \alpha$ .

It follows immediately from Definition 5.2 that if  $\mathcal{I}$  is a collection of independent sets in  $\tau$  and  $\xi_{\mathcal{I}}$  is the corresponding partition, then

$$\xi_{\mathcal{I}}^{(k)} \cap M = \{I_{j_1} \cap T^{-1} I_{j_2} \cap \dots \cap T^{-k+1} I_{j_k} \cap M: j_l \in \{1, \dots, N\}\}$$

is a partition of  $M$  into  $N^k$  sets of equal measure  $\mu(M)/N^k$  for every  $k = 1, \dots, m$ .

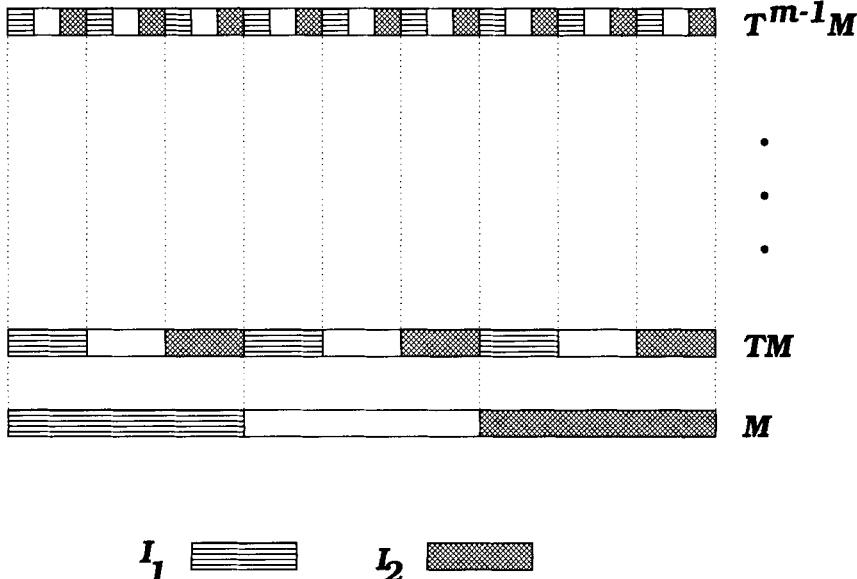


Figure 1. Independent sets  $(I_1, I_2)$  in  $\tau = (M, TM, \dots, T^{m-1}M)$ .

Using (5.2) we easily obtain an estimate on the Rényi entropy of  $\xi_{\mathcal{I}}^{(m)}$ :

$$\begin{aligned} H_{\mu}\left(q, \xi_{\mathcal{I}}^{(m)}\right) &\geq H_{\mu}\left(q, \xi_{\mathcal{I}}^{(m)}|M\right) = \frac{1}{1-q} \log N^m \left(\frac{\mu(M)}{N^m}\right)^q \\ &= m \log N + \frac{q}{1-q} \log \mu(M). \end{aligned}$$

If the measure of the base of the tower  $\mu(M)$  is not too small, say  $\mu(M) \geq N^{-m(1-q)/(2q)}$ , then  $H_{\mu}(q, \xi_{\mathcal{I}}^{(m)})/m \geq (\log N)/2$ .

In the next subsection we estimate the Rényi entropy of a partition which is ‘close’ to some partition  $\xi_{\mathcal{I}}$ , where  $\mathcal{I}$  is a collection of independent sets.

**5.2. APPROXIMATION LEMMA.** Let  $\mathcal{I} = (I_1, \dots, I_{N-1})$  be a collection of independent sets in the tower  $\tau = (M, \dots, T^{m-1}M)$ , and let  $\xi_{\mathcal{I}} = \{I_1, \dots, I_N\}$  be the corresponding partition. Suppose that another partition  $\eta = (E_1, \dots, E_N)$  is such that the sets

$$E_j^{\tau} = E_j \cap \tau$$

are “close” to the corresponding  $I_j$ ’s for  $j = 1, \dots, N - 1$ . Since the partition  $\xi_{\mathcal{I}}^{(m)}$  has a large Rényi entropy (subject to a relation between  $N$  and  $\mu(\tau)$  of course), then the partition  $\eta^{(m)}$  has a large Rényi entropy as well. This can be rigorously formulated in the following way.

LEMMA 5.3: *Let  $\mathcal{I} = \{I_1, \dots, I_{N-1}\}$  be a collection of independent sets in  $\tau = (M, \dots, T^{m-1}M)$  with  $m \geq 16$ , and let  $\xi_{\mathcal{I}} = \{I_1, \dots, I_{N-1}, I_N := X \setminus \bigcup_{j=1}^{N-1} I_j\}$  be the corresponding partition. Suppose  $\eta = (E_1, \dots, E_{N-1}, E_N)$  is another partition of  $X$  such that*

$$(5.3) \quad \sum_{j=1}^{N-1} \mu(E_j^\tau \Delta I_j) \leq \frac{\mu(\tau)}{16(N+31)}.$$

Then for every  $q \in (0, 1)$  one has

$$\frac{1}{m} H_\mu(q, \eta^{(m)}) \geq \frac{1}{4} \log N - \log 2 + \frac{q \log \mu(\tau)}{(1-q)m} - \frac{1}{1-q} \log 2m^q.$$

*Proof:* This lemma is a generalization of lemmas 2.6 and 2.7 from [3], and its proof follows quite closely the proofs of the corresponding results in [3]. Nevertheless, due to the necessary modifications and for the sake of completeness we provide a proof here.

We shall use the following notation: let  $\Omega = \{1, \dots, N\}$  and

$$\begin{aligned} \Delta(\mathbf{r}) &= I_{r_1} \cap T^{-1}I_{r_2} \cap \dots \cap T^{-m+1}I_{r_m} \quad \text{for } \mathbf{r} = (r_1, \dots, r_m) \in \Omega^m, \\ \tilde{\Delta}(\mathbf{s}) &= E_{s_1} \cap T^{-1}E_{s_2} \cap \dots \cap T^{-m+1}E_{s_m} \quad \text{for } \mathbf{s} = (s_1, \dots, s_m) \in \Omega^m. \end{aligned}$$

We know that  $\mu(\Delta(\mathbf{r}) \cap M) = \mu(M)/N^m$ . Since  $\eta$  is close to  $\xi_{\mathcal{I}}$  in  $\tau$ , we expect the sets  $\eta^{(m)} \cap M$  to have approximately the same measure as the sets of  $\xi^{(m)} \cap M$ . Let us make it precise. We say that  $\tilde{\Delta}(\mathbf{s})$ ,  $\mathbf{s} \in \Omega^m$ , is a ‘bad’ (or a ‘fat’) element of  $\eta^{(m)}$  if

$$\mu(\tilde{\Delta}(\mathbf{s}) \cap M) \geq 2^m N^{-m/4} \mu(M),$$

and is ‘good’ (or ‘thin’) otherwise. We collect the indexes of all ‘bad’ elements into the set

$$S = \{\mathbf{s} \in \Omega^m : \tilde{\Delta}(\mathbf{s}) \text{ is ‘bad’}\}.$$

We will now show that ‘bad’ elements of  $\eta^{(m)}$  cover less than a half of  $M$  in measure, i.e.,

$$\mu\left(\bigcup_{\mathbf{s} \in S} \tilde{\Delta}(\mathbf{s}) \cap M\right) \leq \frac{1}{2} \mu(M).$$

We introduce the following notation:

$$\begin{aligned} M(\mathbf{s}, \mathbf{r}) &= \tilde{\Delta}(\mathbf{s}) \cap \Delta(\mathbf{r}) \cap M, \\ \tau(\mathbf{s}, \mathbf{r}) &= \bigcup_{k=0}^{m-1} T^k M(\mathbf{s}, \mathbf{r}), \\ \tau(\mathbf{s}) &= \bigcup_{\mathbf{r} \in \Omega^m} \tau(\mathbf{s}, \mathbf{r}). \end{aligned}$$

It is easy to see that

$$\begin{aligned} M(\mathbf{s}, \mathbf{r}) \cap M(\mathbf{s}, \mathbf{t}) &= \emptyset \quad \text{for } \mathbf{r} \neq \mathbf{t}, \\ T^i M(\mathbf{s}, \mathbf{r}) \cap T^j M(\mathbf{s}, \mathbf{r}) &= \emptyset \quad \text{for } i \neq j, \\ \tau(\mathbf{s}, \mathbf{r}) \cap \tau(\mathbf{s}, \mathbf{t}) &= \emptyset \quad \text{for } \mathbf{r} \neq \mathbf{t}. \end{aligned}$$

Let  $\mathbf{s} \in \Omega^m$ ; then

$$\begin{aligned} (5.4) \quad \mu\left(\left(\bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j\right) \cap \tau(\mathbf{s})\right) &= \sum_{\mathbf{r} \in \Omega^m} \mu\left(\left(\bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j\right) \cap \tau(\mathbf{s}, \mathbf{r})\right) \\ &= \sum_{\mathbf{r} \in \Omega^m} \sum_{k=0}^{m-1} \mu\left(\left(\bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j\right) \cap T^k M(\mathbf{s}, \mathbf{r})\right). \end{aligned}$$

Consider the sets participating in the last sum separately. We claim that

$$(5.5) \quad \left(\bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j\right) \cap T^k M(\mathbf{s}, \mathbf{r}) = \begin{cases} T^k M(\mathbf{s}, \mathbf{r}), & \text{if } s_k \neq r_k, \\ \emptyset, & \text{if } s_k = r_k. \end{cases}$$

For this it is sufficient to show that

$$(5.6) \quad \left(\bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j\right) \cap E_{s_k}^\tau \cap I_{r_k} = \begin{cases} E_{s_k}^\tau \cap I_{r_k} & \text{if } s_k \neq r_k, \\ \emptyset, & \text{if } s_k = r_k. \end{cases}$$

The proof is straightforward: let  $j = 1, \dots, N-1$  and  $k = 1, \dots, m$ ; then

$$\begin{aligned} (E_j^\tau \Delta I_j) \cap E_{s_k}^\tau \cap I_{r_k} &= \left((E_j^\tau \setminus I_j) \cap E_{s_k}^\tau \cap I_{r_k}\right) \cup \left((I_j \setminus E_j^\tau) \cap E_{s_k}^\tau \cap I_{r_k}\right) \\ &=: A \cup B. \end{aligned}$$

Suppose first that  $s_k = r_k$ . Then for  $j = s_k = r_k$  we have

$$A \cup B \subseteq \left((E_j^\tau \setminus I_j) \cap I_j\right) \cup \left((I_j \setminus E_j^\tau) \cap E_j\right) = \emptyset,$$

since  $I_j \subset \tau$  for  $j = 1, \dots, N-1$ .

For  $j \neq s_k = r_k$  we have

$$A \cup B \subseteq \left( (E_j^\tau \setminus I_j) \cap E_{s_k}^\tau \right) \cup \left( (I_j \setminus E_j^\tau) \cap I_{r_k} \right) = \emptyset,$$

since  $E_j \cap E_{s_k} = I_j \cap I_{r_k} = \emptyset$ .

Now consider the case  $s_k \neq r_k$ . If  $j \neq s_k$  and  $j \neq r_k$ , then

$$A \cup B \subseteq (E_j \cap E_{s_k}) \cup (I_j \cap I_{r_k}) = \emptyset.$$

If  $j \neq s_k$  and  $j = r_k$ , then  $A \subseteq (E_j^\tau \setminus I_j) \cap E_{s_k} = \emptyset$ , but

$$B = (I_j \setminus E_j^\tau) \cap E_{s_k} \cap I_{r_k} = (I_{r_k} \cap E_{s_k}^\tau),$$

since  $E_j^\tau \cap E_{s_k}^\tau = \emptyset$ .

Similarly, for  $j = s_k$  and  $j \neq r_k$ , we conclude that  $B = \emptyset$ , but  $A = E_{s_k}^\tau \cap I_{r_k}$ . Hence we proved (5.6), and therefore (5.5).

Using (5.5) and the fact that  $T$  is measure-preserving, we can simplify (5.4):

$$(5.7) \quad \mu \left( \left( \bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j \right) \cap \tau(\mathbf{s}) \right) = \sum_{\mathbf{r} \in \Omega^m} d_H(\mathbf{s}, \mathbf{r}) \mu(M(\mathbf{s}, \mathbf{r})),$$

where  $d_H(\mathbf{s}, \mathbf{r}) = \#\{k: s_k \neq r_k\}$  is the **Hamming distance** between  $\mathbf{s}$  and  $\mathbf{r}$ . We rewrite (5.7) in the following form:

$$(5.8) \quad \mu \left( \left( \bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j \right) \cap \tau(\mathbf{s}) \right) = \sum_{i=0}^m i \sum_{\mathbf{r}: d_H(\mathbf{s}, \mathbf{r})=i} \mu(M(\mathbf{s}, \mathbf{r})).$$

Given  $\mathbf{s} \in \Omega^m$ , the number of  $\mathbf{r}$ 's such that  $d_H(\mathbf{s}, \mathbf{r}) = i$  is  $C_m^i (N-1)^i$ , where  $C_m^i$  is the binomial coefficient. Let us introduce the following notation:

$$x_i(\mathbf{s}) := \sum_{\mathbf{r}: d_H(\mathbf{s}, \mathbf{r})=i} \mu(M(\mathbf{s}, \mathbf{r})), \quad y_i := \frac{\mu(M)}{N^m} C_m^i (N-1)^i.$$

Note that  $\mu(\tilde{\Delta}(\mathbf{s}) \cap M) = \sum_{i=0}^m x_i(\mathbf{s})$ .

Since  $M(\mathbf{s}, \mathbf{r}) \subset \Delta(\mathbf{r}) \cap M$  and  $\mu(\Delta(\mathbf{r}) \cap M) = \mu(M)/N^m$ , for every  $i$  one has

$$(5.9) \quad \begin{aligned} x_i(\mathbf{s}) &= \sum_{\mathbf{r}: d_H(\mathbf{s}, \mathbf{r})=i} \mu(M(\mathbf{s}, \mathbf{r})) \leq \frac{\mu(M)}{N^m} \left[ \sum_{\mathbf{r}: d_H(\mathbf{s}, \mathbf{r})=i} 1 \right] \\ &= \frac{\mu(M)}{N^m} C_m^i (N-1)^i = y_i. \end{aligned}$$

Furthermore, for every  $\mathbf{s}$  there exists  $k_{\mathbf{s}} \in \{1, \dots, m\}$  such that

$$(5.10) \quad \sum_{i=0}^{k_{\mathbf{s}}} y_i > \sum_{i=0}^m x_i(\mathbf{s}) \geq \sum_{i=0}^{k_{\mathbf{s}}-1} y_i.$$

From (5.9) and (5.10) we conclude that  $\sum_{i=l}^m x_i(\mathbf{s}) \geq \sum_{i=l}^{k_{\mathbf{s}}-1} y_i$  for all  $l \geq 0$ , and as a result

$$(5.11) \quad \sum_{i=0}^m i x_i(\mathbf{s}) \geq \sum_{i=0}^{k_{\mathbf{s}}-1} i y_i.$$

We will show now that if  $\mathbf{s} \in S$  then  $k_{\mathbf{s}} \geq [3m/4] + 1$ . Indeed, if  $\mathbf{s} \in S$ , then by the definition of  $S$ ,

$$\mu(\tilde{\Delta}(\mathbf{s}) \cap M) \geq 2^m N^{-m/4} \mu(M),$$

and from (5.9) we have

$$\frac{1}{N^m} \sum_{i=0}^{k_{\mathbf{s}}} C_m^i (N-1)^i \geq \frac{\sum_{i=0}^m x_i(\mathbf{s})}{\mu(M)} = \frac{\mu(\tilde{\Delta}(\mathbf{s}) \cap M)}{\mu(M)} \geq 2^m N^{-m/4}.$$

However, Lemma 6.1 (see Appendix below) states that

$$\frac{1}{N^m} \sum_{i=0}^k C_m^i (N-1)^i < 2^m N^{-m/4}$$

for all  $k = 0, 1, \dots, [3m/4]$ . Hence,  $k_{\mathbf{s}} \geq [3m/4] + 1$ .

Now, for all  $\mathbf{s} \in S$  we have

$$\begin{aligned} \mu\left(\left(\bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j\right) \cap \tau(\mathbf{s})\right) &= \sum_{i=0}^m i x_i(\mathbf{s}) \quad (\text{by (5.8)}) \\ &\geq \sum_{i=0}^{k_{\mathbf{s}}-1} i y_i = \frac{\mu(M)}{N^m} \sum_{i=0}^{k_{\mathbf{s}}-1} i C_m^i (N-1)^i \quad (\text{by (5.11)}) \\ &\geq \frac{\mu(M)}{N^m} \frac{m}{8(N+31)} \sum_{i=0}^{k_{\mathbf{s}}} C_m^i (N-1)^i \quad (\text{by Lemma 6.1}) \\ &= \frac{m}{8(N+31)} \sum_{i=0}^{k_{\mathbf{s}}} y_i \\ &\geq \frac{m}{8(N+31)} \sum_{i=0}^m x_i(\mathbf{s}) \quad (\text{by (5.10)}). \end{aligned}$$

Hence,

$$\begin{aligned}
\mu\left(\bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j\right) &\geq \sum_{\mathbf{s} \in S} \mu\left(\left(\bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j\right) \cap \tau(\mathbf{s})\right) \\
&\geq \frac{m}{8(N+31)} \sum_{\mathbf{s} \in S} \sum_{i=0}^m x_i(\mathbf{s}) = \frac{m}{8(N+31)} \sum_{\mathbf{s} \in S} \sum_{\mathbf{r} \in \Omega^m} \mu(M(\mathbf{s}, \mathbf{r})) \\
&= \frac{m}{8(N+31)} \sum_{\mathbf{s} \in S} \mu(\tilde{\Delta}(\mathbf{s}) \cap M).
\end{aligned}$$

Therefore, using our assumption (5.3) one has

$$\begin{aligned}
\sum_{\mathbf{s} \in S} \mu(\tilde{\Delta}(\mathbf{s}) \cap M) &\leq \frac{8(N+31)}{m} \mu\left(\bigcup_{j=1}^{N-1} E_j^\tau \Delta I_j\right) \\
&\leq \frac{8(N+31)}{16(N+31)m} \mu(\tau) = \frac{1}{2} \mu(M),
\end{aligned}$$

i.e., ‘bad’ elements  $\tilde{\Delta}(\mathbf{s})$  cover not more than a half of  $M$ .

Now we can estimate the Rényi entropy of the partition  $\eta^{(m)}$ :

$$\begin{aligned}
H_\mu(q, \eta^{(m)}) &\geq H_\mu(q, \eta^{(m)}|M) = \frac{1}{1-q} \log \left( \sum_{\mathbf{s} \in \Omega^m} \mu(\tilde{\Delta}(\mathbf{s}) \cap M)^q \right) \\
&\geq \frac{1}{1-q} \log \left( \sum_{\mathbf{s} \in \Omega^m \setminus S} \mu(\tilde{\Delta}(\mathbf{s}) \cap M)^q \right) \\
&\geq \frac{1}{1-q} \log \left( \sum_{\mathbf{s} \in \Omega^m \setminus S} \frac{\mu(\tilde{\Delta}(\mathbf{s}) \cap M)}{(2^m N^{-m/4} \mu(M))^{1-q}} \right) \\
&\geq \frac{1}{1-q} \log \left( \frac{1}{2} \frac{\mu(M)}{(2^m N^{-m/4} \mu(M))^{1-q}} \right) \\
&= \frac{m}{4} \log N - m \log 2 + \frac{q}{1-q} \log \mu(M) - \frac{1}{1-q} \log 2 \\
&= m \left( \frac{1}{4} \log N - \log 2 + \frac{q \log \mu(\tau)}{(1-q)m} \right) - \frac{1}{1-q} \log 2 m^q.
\end{aligned}$$

This finishes the proof of Lemma 5.3. ■

**5.3. PARTITIONS WITH LARGE RÉNYI ENTROPY.** Consider  $q \in (0, 1)$  and take  $N \in \mathbb{N}$ ,  $N \geq 16$ . For the convenience of notation we put  $\delta = (1-q)/8q$ . Take  $R \in \mathbb{N}$  such that

$$N^{\delta R} > 32(N+31)N.$$

We choose a sequence of Rokhlin–Halmos towers  $\{\tau_k\}$ ,

$$\tau_k = (M_k, \dots, T^{m_k-1}M_k)$$

of height  $m_k = Rk$  and measure  $\mu(\tau_k) = N^{-\delta Rk}$ . For each  $k$  let

$$\mathcal{I}_k = (I_1(k), \dots, I_{N-1}(k))$$

be a collection of independent sets in  $\tau_k$ . We define a sequence of collections of pairwise disjoint sets  $\mathcal{E}_k = (E_1(k), \dots, E_{N-1}(k))$  as follows: for  $j = 1, \dots, N-1$  let

$$E_j(0) = \emptyset,$$

$$E_j(k) = (E_j(k-1) \setminus \tau_k) \cup I_j(k) \quad \text{for } k = 1, 2, \dots$$

For any  $j \in \{1, \dots, N-1\}$  the sequence of characteristic functions  $\{\chi_{E_j(k)}\}_{k=1}^\infty$  is a Cauchy sequence in  $L_1(X, \mathfrak{B}, \mu)$ . Indeed, we obviously have

$$E_j(k) \Delta E_j(k-1) \subseteq \tau_k,$$

and hence for  $k_1, k_2 \geq K$  we have

$$\mu(E_j(k_1) \Delta E_j(k_2)) \leq \sum_{k=K}^{+\infty} \mu(\tau_k) \rightarrow 0 \quad \text{as } K \rightarrow \infty.$$

From this we conclude that there exists  $E_j \in \mathfrak{B}$  such that

$$\chi_{E_j(k)} \rightarrow \chi_{E_j} \quad \text{for } k \rightarrow \infty.$$

It follows from the construction that  $\mu(E_j \cap E_i) = 0$  for  $i \neq j$ . Since we can neglect sets of measure zero we may assume that  $E_i \cap E_j = \emptyset$  and hence we have a collection  $\mathcal{E} = (E_1, \dots, E_{N-1})$  of pairwise disjoint subsets of  $X$ .

Furthermore, for every  $j = 1, \dots, N-1$  and any  $L > k$  one has

$$\begin{aligned} \mu(E_j^{\tau_k} \Delta I_j(k)) &\leq \mu(E_j^{\tau_k} \Delta E_j^{\tau_k}(L)) + \sum_{l=k}^{L-1} \mu(E_j^{\tau_k}(l+1) \Delta E_j^{\tau_k}(l)) \\ &\quad + \mu(E_j^{\tau_k}(k) \Delta I_j(k)). \end{aligned}$$

Moreover, since  $E_j^{\tau_k}(k) = I_j(k)$ ,  $\mu(E_j \Delta E_j(L)) \rightarrow 0$  as  $L \rightarrow \infty$ , and  $E_j(l+1) \Delta E_j(l) \subseteq \tau_l$ , we conclude that

$$\mu(E_j^{\tau_k} \Delta I_j(k)) \leq \sum_{l=k+1}^{\infty} \mu(\tau_l) = \mu(\tau_k) \frac{N^{-\delta R}}{1 - N^{-\delta R}} < \frac{\mu(\tau_k)}{16(N+31)N},$$

and hence

$$\sum_{j=1}^{N-1} \mu(E_j^{\tau_k} \Delta I_j(k)) < \frac{\mu(\tau_k)}{16(N+31)}.$$

Now let  $\eta = \{E_1, \dots, E_{N-1}, E_N\}$ , where  $E_N = X \setminus \bigcup_{j=1}^{N-1} E_j$ , and applying Lemma 5.3 we conclude that

$$\begin{aligned} \frac{1}{m_k} H_\mu(q, \eta^{(m_k)}) &\geq \frac{1}{4} \log N + \frac{q}{(1-q)m_k} \log \mu(\tau_k) + o(1) \\ &= \frac{1}{8} \log N + o(1), \end{aligned}$$

where  $o(1) \rightarrow 0$  as  $k \rightarrow \infty$ .

For any  $n \in \mathbb{N}$  there exists  $k \in \mathbb{N}$  such that

$$m_k = Rk \leq n < R(k+1) = m_{k+1}.$$

Since  $H_\mu(q, \eta^{(n)}) \geq H_\mu(q, \eta^{(m_k)})$ , we have

$$\frac{1}{n} H_\mu(q, \eta^{(n)}) \geq \frac{m_k}{m_{k+1}} \frac{1}{m_k} H_\mu(q, \eta^{(m_k)}) \geq \frac{1}{1 + \frac{1}{k}} \left( \frac{1}{8} \log N + o(1) \right).$$

This proves that

$$h_\mu(T, q, \eta) = \liminf_{n \rightarrow \infty} \frac{1}{n} H_\mu(q, \eta^{(n)}) \geq \frac{1}{9} \log N.$$

Everywhere above we have assumed that  $q \in (0, 1)$ . However, since

$$h(T, \mu, q, \xi) \geq h(T, \mu, 1/2, \xi)$$

for all  $q \leq 0$  and every partition  $\xi$ , we have obtained partitions with large Rényi entropies of all orders  $q$ ,  $q \leq 0$ , as well. Finally, since  $N$  is an arbitrary integer, we proved the remaining part of Theorem 1.2.

## 6. Final remarks

(a) Another version of Rényi entropies can be defined using  $\limsup$  instead of  $\liminf$  in (2.2). In principle, due to the lack of subadditivity of  $H_q$ , there might exist a finite partition  $\xi$  such that

$$\begin{aligned} \bar{h}(T, \mu, q, \xi) &:= \limsup_{n \rightarrow \infty} -\frac{1}{(q-1)n} H_\mu(q, \xi^{(n)}) \\ &\neq \liminf_{n \rightarrow \infty} -\frac{1}{(q-1)n} H_\mu(q, \xi^{(n)}) = h(T, \mu, q, \xi) \end{aligned}$$

for some  $q \in \mathbb{R}$ ,  $q \neq 1$ . However, using the results of Theorems 1.1 and 1.2, one can easily show that

$$\bar{h}(T, \mu, q) := \sup_{\xi \text{ is finite}} \bar{h}(T, \mu, q, \xi) = \sup_{\xi \text{ is finite}} h(T, \mu, q, \xi) = h(T, \mu, q)$$

for all  $q \in \mathbb{R}$ .

Since  $h(T, \mu, q) = +\infty$  for  $q < 1$  in the ergodic and aperiodic cases, the claim is obviously true for  $q < 1$ .

To complete the proof we have to show that, for  $q > 1$ ,

$$(6.1) \quad \sup_{\xi \text{ is finite}} \bar{h}(T, \mu, q, \xi) \leq h(T, \mu)$$

in the ergodic case, and

$$(6.2) \quad \sup_{\xi \text{ is finite}} \bar{h}(T, \mu, q, \xi) \leq h_*(T, \mu)$$

in the aperiodic case. The first inequality (6.1) follows immediately from the monotonicity properties (Proposition 2.1), and the fact that for  $q = 1$  (standard entropy) the limit in (2.2) exists.

The second inequality (6.2) is proved exactly in the same manner as an inequality  $h(T, \mu, q) \leq h_*(T, \mu)$  in section 4.1.

(b) Formally, the pair of metric invariants  $(h(T, \mu), h(T, \mu, q))$ ,  $q > 1$ , can detect ergodicity: if  $h(T, \mu) - h(T, \mu, q) > 0$ , then  $(T, \mu)$  cannot be ergodic. However, we were not able to find any relevant examples where this could be useful.

In our opinion, an example of a system  $(X, \mathfrak{B}, \mu, T)$ , where the non-ergodicity can be decided from the positiveness of  $h(T, \mu) - h(T, \mu, q)$ , would be interesting.

(c) The difference between ergodicity and non-ergodicity is less interesting than the difference between ergodicity and weak mixing. As is well known, weak mixing of  $T$  is equivalent to the ergodicity of any direct products of  $T$  with an ergodic automorphism  $S$ . Suppose  $T$  is ergodic, but not weakly mixing. Then there exists an ergodic measure-preserving dynamical system  $(Y, \mathfrak{F}, \nu, S)$  such that  $(X \times Y, \mathfrak{G}, \mu \times \nu, T \times S)$  is not ergodic. Unfortunately, the Rényi entropies are not able to detect non-ergodicity of such systems: for  $q > 1$  one has

$$h(T \times S, \mu \times \nu, q) = h(T, \mu, q) + h(S, \nu, q) = h(T, \mu) + h(S, \nu) = h(T \times S, \mu \times \nu)$$

where first and the third equalities are standard facts for entropy-like characteristics, and the second equality follows from Theorem 1.1

(d) *Entropy convergence rates* were introduced in [3].

Let  $(X, \mathfrak{B}, \mu)$  be a Lebesgue space and  $T$  be a measure-preserving automorphism. Suppose that  $(X, \mathfrak{B}, \mu, T)$  has zero entropy. Hence, for any finite partition  $\xi$  one has

$$h(T, \mu, \xi) = \liminf_{n \rightarrow \infty} \frac{1}{n} H_\mu(\xi^{(n)}) = 0.$$

Let  $c > 0$  and  $a_n$ ,  $n \geq 1$ , is a sequence of positive numbers such that  $a_n \rightarrow \infty$ . Denote by  $\Pi$  the set of all non-trivial partitions  $(X, \mathfrak{B}, \mu)$  into two sets.

The automorphism  $T$  is said to be

- of type  $(LI \geq c)$  for  $((a_n), \Pi)$  if, for every  $\xi \in \Pi$ ,

$$\liminf_{n \rightarrow \infty} \frac{1}{a_n} H(\xi^{(n)}) \geq c;$$

- of type  $(LS \geq c)$  for  $((a_n), \Pi)$  if, for every  $\xi \in \Pi$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{a_n} H(\xi^{(n)}) \geq c.$$

Similarly one defines types  $(LI \leq c)$ ,  $(LS \leq c)$ ,  $(LI < c)$ , etc. Clearly, the type of a transformation is a measure-theoretic invariant.

It was shown in [3] that there are no aperiodic transformations of type  $(LI < \infty)$  for  $((a_n), \Pi)$ , where  $a_n = o(n)$ ,  $n \geq 1$ . Every totally ergodic transformation (i.e.,  $T^k$  is ergodic for every  $k \geq 1$ ) is of type  $(LS = \infty)$  for  $(g(\log n), \Pi)$ , where  $g: [0, +\infty) \rightarrow \mathbf{R}$  is positive, monotone increasing and

$$\int_1^\infty \frac{g(x)}{x^2} dx < \infty.$$

Also, in [2] F. Blume constructed a class of weakly mixing systems, which can be distinguished by these invariants.

It would be interesting to know if the corresponding notions for the Rényi entropies, in the case of both  $q < 1$  and  $q > 1$ , can produce useful convergence rates, which are different from the case of standard entropy.

## Appendix. Auxiliary results

Throughout this section we assume that  $N \in \mathbb{N}$ ,  $N \geq 2$ ,  $[x]$  denotes an integer part of  $x$ , and  $C_m^k$  will denote the binomial coefficient

$$C_m^k = \frac{m!}{k!(m-k)!}.$$

For  $k = 0, \dots, m$  let

$$a_m(k, N) = \sum_{i=0}^k C_m^i (N-1)^i, \quad b_m(k, N) = \sum_{i=0}^{k-1} i C_m^i (N-1)^i.$$

The following result is a straightforward generalization of Lemma 2.4 from [3].

LEMMA 6.1: *Let  $m$  be an integer,  $m \geq 16$ , and put  $k_0 = [3m/4] + 1$ . Then*

(1) *for  $k = 0, 1, \dots, k_0 - 1$  one has*

$$(6.3) \quad \frac{a_m(k, N)}{N^m} \leq 2^m N^{-m/4};$$

(2) *for  $k = k_0, \dots, m$  one has*

$$(6.4) \quad \frac{a_m(k, N)}{b_m(k, N)} \leq \frac{8(N + 31)}{m}.$$

ACKNOWLEDGEMENT: We are grateful to F. Blume and E. G. F. Thomas for valuable discussions. The second author acknowledges the support of the Dutch Organization for Scientific Research (NWO), grant 613-06-551.

## References

- [1] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems*, Cambridge University Press, Cambridge, 1993.
- [2] F. Blume, *Fast entropy convergence for rank one systems*, Technical report, 1997.
- [3] F. Blume, *Possible rates of entropy convergence*, Ergodic Theory and Dynamical Systems **17** (1997), 45–70.
- [4] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic Theory* (translated from the Russian by A. B. Sosinskiĭ), Springer-Verlag, New York, 1982.
- [5] M. Denker, C. Grillenberger and K. Sigmund, *Ergodic Theory on Compact Spaces*, Lecture Notes in Mathematics, Vol. 527, Springer-Verlag, Berlin, 1976.
- [6] J.-P. Eckmann and D. Ruelle, *Ergodic theory of chaos and strange attractors*, Reviews of Modern Physics **57** (1985), 617–656.
- [7] P. Grassberger and I. Procaccia, *Dimensions and entropies of strange attractors from a fluctuating dynamics approach*, Physica D **13** (1984), 34–54.
- [8] H. G. E. Hentschel and I. Procaccia, *The infinite number of generalized dimensions of fractals and strange attractors*, Physica D **8** (1983), 435–444.
- [9] J. C. Kieffer, *On the minimum number of fixed length sequences with fixed total probability*, Annals of Probability **4** (1976), 335–337.
- [10] J. C. Kieffer and M. Rahe, *Selecting universal partitions in ergodic theory*, The Annals of Probability **9** (1981), 705–709.
- [11] W. Krieger, *On entropy and generators of measure-preserving transformations*, Transactions of the American Mathematical Society **149** (1970), 453–464.

- [12] D. S. Ornstein, *Ergodic Theory, Randomness, and Dynamical Systems*, James K. Whittemore Lectures in Mathematics given at Yale University, Yale Mathematical Monographs, No. 5, Yale University Press, New Haven, Conn., 1974.
- [13] Ya. B. Pesin, *Dimension Theory in Dynamical Systems*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 1997.
- [14] A. Rényi, *Probability Theory*, translated by László Vekerdi, North-Holland Series in Applied Mathematics and Mechanics, Vol. 10, North-Holland, Amsterdam, 1970.
- [15] V. A. Rohlin, *On the fundamental ideas of measure theory*, Matematicheskii Sbornik (N.S.) **25**(67) (1949), 107–150.
- [16] V. A. Rokhlin, *Izbrannye raboty*, reminiscences of V. A. Rokhlin by V. I. Arnol'd, A. M. Vershik, S. P. Novikov and Ya. G. Sinai, edited and with a preface by A. M. Vershik, Moskovskii Tsentr Nepreryvnogo Matematicheskogo Obrazovaniya, Moscow, 1999, p. 496.
- [17] D. J. Rudolph, *Fundamentals of Measurable Dynamics*, Ergodic theory on Lebesgue spaces, The Clarendon Press, Oxford University Press, New York, 1990.
- [18] P. Shields, *The Theory of Bernoulli Shifts*, Chicago Lectures in Mathematics, The University of Chicago Press, Chicago, 1973.
- [19] Ja. G. Sinaĭ, *On a weak isomorphism of transformations with invariant measure*, Matematicheskii Sbornik (N.S.) **63**(105) (1964), 23–42.
- [20] M. Smorodinsky, *Ergodic Theory, Entropy*, Lecture Notes in Mathematics, Vol. 214, Springer-Verlag, Berlin, 1971.
- [21] Š. Šujan, *Generators of an abelian group of invertible measure-preserving transformations*, Monatshefte für Mathematik **90** (1980), 67–79.
- [22] F. Takens and E. Verbitskiy, *Generalized entropies: Rényi and correlation integral approach*, Nonlinearity **11** (1998), 771–782.
- [23] K. Winkelbauer, *On the asymptotic rate of non-ergodic information sources*, Kybernetika (Prague) **6** (1970), 127–148.