Ergod. Th. & Dynam. Sys. (1997), **17**, 45–70
Printed in the United Kingdom © *1997 Cambridge University Press*

Possible rates of entropy convergence

FRANK BLUME
Department of Mathematics, John Brown University, Siloam Springs, AR 72761, USA
(e-mail: fblume@acc.jbu.edu)

(*Received* 20 *January* 1995 *and revised* 15 *October* 1995)

Abstract. We introduce entropy convergence rates as isomorphism invariants for measure-preserving systems and prove several general facts concerning these rates for aperiodic systems, completely ergodic systems and rank-one systems. We will for example show that for any completely ergodic system (X, T) and any non-trivial partition α of X into two sets we have $\limsup_{n\to\infty} H(\alpha_0^{n-1})/g(\log_2 n) = \infty$, whenever g is a positive increasing function on $(0, \infty)$ such that $g(x)/x^2$ is integrable.

0. *Introduction*

The concept of entropy convergence rates is designed to provide new isomorphism invariants which can be used to distinguish certain measure-preserving systems with entropy zero. It is conceivable that they can also be used to distinguish systems with equal positive entropy, but this is an open question and not a subject of our discussion. We focus our attention exclusively on a particular type of entropy convergence rate which is defined via the standard definition for the entropy of a finite partition. §5 presents a variety of related concepts that can be used to design similar isomorphism invariants. There seems to be a close connection to the invariants of Kakutani equivalence which were used by Ratner in [**7**] to show that different cartesian powers of the horocycle flow are inequivalent and possibly there is also a relation to other convergence type problems as discussed by Shields in [**10**].

That entropy convergence rates can be used to distinguish systems up to isomorphism is shown in [**2**]. The content of the present paper is part of the author's doctoral dissertation [**1**] which can serve as a reference for further details of some of the proofs given here.

If $(X, \mathcal{B}, \mu)$ is a probability space, $T : X \to X$ a measure-preserving transformation and α is a finite partition of X, then we denote by α_0^{n-1} the nth refinement of α under T. Furthermore, if $f(x) := -x \log_2 x$ for $x \in [0, 1]$, then the *entropy of a partition* α is defined as

$$H(\alpha) := \sum_{A\in\alpha} f(\mu(A)).$$

We will now introduce the different types of entropy convergence rates. Let (X, T) be a measure-preserving system, let $\{a_n\}_{n=1}^{\infty}$ be a monotone increasing sequence with

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

$\lim_{n\to\infty} a_n = \infty$ and $c \in (0, \infty)$. If P is a class of partitions of X, then we say that (X, T) is of type $(LS \geq c)$ for $(\{a_n\}_{n=1}^{\infty}, P)$ if

$$\limsup_{n\to\infty} \frac{H(\alpha_0^{n-1})}{a_n} \geq c \quad \text{for all } \alpha \in P$$

and (X, T) is of type $(LI \geq c)$ for $(\{a_n\}_{n=1}^{\infty}, P)$ if

$$\liminf_{n\to\infty} \frac{H(\alpha_0^{n-1})}{a_n} \geq c \quad \text{for all } \alpha \in P.$$

Completely analogously we define the types $(LS \leq c)$, $(LI \leq c)$ and also the special types $(LS < \infty)$, $(LS = \infty)$, $(LI < \infty)$, $(LI = \infty)$, $(LS > 0)$ and $(LI > 0)$. It is trivial to see that if (X, T) is of type $(LI \geq c)$, then it is also of type $(LS \geq c)$ and similarly

$$(LS \leq c) \Rightarrow (LI \leq c), \quad (LS < \infty) \Rightarrow (LI < \infty) \quad \text{and} \quad (LI > 0) \Rightarrow (LS > 0).$$

In §5 we will discuss alternative definitions for entropy convergence types that arise if we use entropy functions other than $f(x) = -x \log_2 x$.

It is a valid question to ask how the convergence type of a measure-preserving system depends on the choice of P, but this will not be part of our discussion. Instead we will *mainly* consider the standard class

$$P(X) := \{\alpha \mid \alpha = \{E, X \setminus E\} \text{ for some } E \in \mathcal{B} \text{ with } 0 < \mu(E) < 1\}$$

and concentrate our attention on the choice of the sequence $\{a_n\}_{n=1}^{\infty}$. The reason for choosing $P(X)$ as our standard class is twofold: on the one hand $P(X)$ is simple enough to reduce the complexity of many proofs, and on the other hand it is large enough for the different convergence types to become isomorphism invariants. More precisely, if (X, T) and (Y, S) are isomorphic measure-preserving systems, then (X, T) is of type $(LS \geq c)$ for $(\{a_n\}_{n=1}^{\infty}, P(X))$ if and only if (Y, S) is of type $(LS \geq c)$ for $(\{a_n\}_{n=1}^{\infty}, P(Y))$. The proof of this statement is very simple and based on the observation that if $\varphi : X \to Y$ is an isomorphism for (X, T) and (Y, S), then

$$H(\alpha_0^{n-1}) = H(\varphi(\alpha)_0^{n-1}) \quad \text{for all } \alpha \in P(X) \text{ and all } n \in \mathbb{N},$$

where $\varphi(\alpha) := \{\varphi(A) \mid A \in \alpha\}$. Analogous statements are of course true for all the other convergence types.

If (X, T) is a measure-preserving system, then $\lim_{n\to\infty} H(\alpha_0^{n-1})/n$ exists and is finite for every finite partition α of X. This is a fundamental result of the theory of entropy, and it shows that every measure-preserving system is of type $(LS < \infty)$ and $(LI < \infty)$ for $(\{n\}_{n=1}^{\infty}, P(X))$. If (X, T) is a *K-system*, then

$$\lim_{n\to\infty} \frac{H(\alpha_0^{n-1})}{n} > 0$$

for every non-trivial partition α of X, and it follows that every K-system is of type $(LS < \infty)$, $(LI < \infty)$, $(LS > 0)$ and $(LI > 0)$ for $(\{n\}_{n=1}^{\infty}, P(X))$. In the case that (X, T) has *entropy zero*, we have

$$\lim_{n\to\infty} \frac{H(\alpha_0^{n-1})}{n} = 0$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

for all finite partitions α of X and, therefore, it is reasonable to consider for systems with entropy zero only convergence types for which

$$\lim_{n\to\infty} \frac{a_n}{n} = 0.$$

Under this assumption we will show in §2 that there are no aperiodic systems of type $(LI < \infty)$ for $(\{a_n\}_{n=1}^{\infty}, P(X))$ and, therefore, there are also none of type $(LS < \infty)$, $(LS \leq c)$ or $(LI \leq c)$.

Furthermore, we will see in §3 that every *completely ergodic* system (i.e. T^n is ergodic for all $n \in \mathbb{Z}$) is of type $(LS = \infty)$ for $(\{g(\log_2 n)\}_{n=1}^{\infty}, P(X))$, whenever $g : [0, \infty) \to \mathbb{R}$ is a positive monotone increasing function with

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

This result will actually be a consequence of a more general theorem for aperiodic systems. Moreover, the constructions in [**2**] will show that the condition $\int_1^{\infty} g(x)/x^2\, dx < \infty$ is the best possible one.

In order to find zero-entropy systems of type $(LS > 0)$ or $(LS \geq c)$ with convergence rates that are faster increasing than $\{g(\log_2 n)\}_{n=1}^{\infty}$ with $\int_1^{\infty} g(x)/x^2\, dx < \infty$, we have to impose stronger conditions on the system (X, T). A result of this type will be mentioned in §4.

Another range of questions concerns the concept of pointwise convergence rates. For this matter we need to set up some more notation. If α is a finite partition of X and $x \in X$, then we denote by $\alpha(x)$ the atom of α to which x belongs and we define the *information function* of α as

$$I_\alpha(x) := -\log_2 \mu(\alpha(x)).$$

It is easy to see that

$$H(\alpha) = \int_X I_\alpha(x)\, d\mu$$

and therefore

$$H(\alpha_0^{n-1}) = \int_X I_{\alpha_0^{n-1}}(x)\, d\mu.$$

This allows us to define pointwise convergence rates as follows. (X, T) is said to be pointwise of type $(LS \geq c)$ for $(\{a_n\}_{n=1}^{\infty}, P)$ if

$$\limsup_{n\to\infty} \frac{I_{\alpha_0^{n-1}}(x)}{a_n} \geq c \text{ a.e.} \quad \text{for all } \alpha \in P.$$

Similarly, all the other convergence types can be defined pointwise a.e. One answer to this type of question will be given in §3, where we show that every completely ergodic system is pointwise of type $(LS \geq 1)$ for $(\{\log_2 n\}_{n=1}^{\infty}, P(X))$. This result will be obtained as a corollary to a more general statement about aperiodic systems. Another result in this context can be obtained from the Shannon–McMillan–Breiman theorem, which says that $I_{\alpha_0^{n-1}}(x)/n$ converges pointwise a.e. to the entropy of T with respect to α. This shows that every measure-preserving system is pointwise of type $(LS < \infty)$ and

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

$(LI < \infty)$ for $(\{n\}_{n=1}^{\infty}, P(X))$ and every K-system is pointwise of type $(LS > 0)$ and $(LI > 0)$ for $(\{n\}_{n=1}^{\infty}, P(X))$. Apart from these few results, the concept of pointwise convergence rates is completely open for further investigation.

1. *Periods of names along towers*

We will introduce some notation and prove several rather technical lemmas which will be needed throughout the rest of our discussion.

In order to show that (X, T) is of a certain convergence type as described in the introduction, we need to find estimates for $H(\alpha_0^{n-1})$ if $\alpha \in P(X)$ and this requires that we find ways to analyze the refined partitions α_0^{n-1}. Our most important tool will be the symbolic representation of atoms in α_0^{n-1} by their 01-names as explained in Definition 1.3. This symbolic representation will be combined with the use of Rokhlin towers $\tau = (M, TM, \ldots, T^{n-1}M)$ which will cover a sufficiently large portion of X. In particular, we will only be interested in the entropy of α_0^{n-1} with respect to the set $\tau = \bigcup_{k=0}^{n-1} T^k M$. (Notice that we use the same symbol for the tower $(M, TM, \ldots, T^{n-1}M)$ and for the set $\bigcup_{k=0}^{n-1} T^k M$ in order to simplify our notation.) This leads us to the following definition.

Definition 1.1. If β is a finite partition of X and $F \in \mathcal{B}$, then we define the entropy of β restricted to F as

$$H_F(\beta) := \sum_{B \in \beta} f(\mu(B \cap F)).$$

The following lemma is a trivial consequence of the fact that $f(x) < 1$ for all $x \in [0, 1]$.

LEMMA 1.2. *If β is a finite partition of X and $G, F \in \mathcal{B}$ with $G \subset F$, then*

$$H_F(\beta) > H_G(\beta) - 2.$$

The most important case for applying this lemma will be when $F = X$, $G = \tau$ and $\beta = \alpha_0^{n-1}$. Then we have $H(\alpha_0^{n-1}) = H_F(\alpha_0^{n-1}) > H_\tau(\alpha_0^{n-1}) - 2$. The following definition gives a precise description of what it means to represent atoms in α_0^{n-1} by 01-names.

Definition 1.3. For any $E \in \mathcal{B}$, $x \in X$, $n \in \mathbb{N}$ and $s \in \{0, 1\}^n$ we define

$$\alpha(E) := \{E, X \setminus E\},$$
$$s_n^E(x) := (\chi_E(x), \chi_E(Tx), \ldots, \chi_E(T^{n-1}x)) \quad \text{and}$$
$$A_n^E(s) := (s_n^E)^{-1}(\{s\}).$$

With these definitions it is easy to see that $\alpha(E)_0^{n-1} = \{A_n^E(s) \mid s \in \{0, 1\}^n\}$. In order to understand how the measure of $A_n^E(s)$ depends on the properties of s, we need to define the period of a 01-name.

Definition 1.4. For $s = (s_0, \ldots, s_{n-1}) \in \{0, 1\}^n$ we define the *period* of s as

$$p_n(s) := \min\{k \in \{1, \ldots, n-1\} \mid s_i = s_{i+k} \text{ for all } i \in \{0, \ldots, n-k-1\}\} \cup \{n\}.$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

This definition says that in order to determine the period of s, we have to find out how far we need to shift s relative to itself to the left to get identical overlapping parts. Here is an example. If $s = (0, 0, 1, 0, 0, 1, 0, 0) \in \{0, 1\}^8$, then we need to shift s by 3 positions to the left to obtain identical overlapping parts:

$$00100100$$
$$00100100.$$

Therefore, we have $p_8((0, 0, 1, 0, 0, 1, 0, 0)) = 3$.

Now let $x \in X$, $E \in \mathcal{B}$ and $s \in \{0, 1\}^n$. We ask the following question: how does the maximum frequency of occurrence of s along the orbit of x depend on the period of s? More precisely: what is the maximum number of $k \in \{0, \ldots, n-1\}$ for which $s_n^E(T^k x) = s$? The following answer is easily obtained.

LEMMA 1.5. *Let $x \in X$, $E \in \mathcal{B}$ and $s \in \{0, 1\}^n$. Then*

$$\#\{k \in \{0, \ldots, n-1\} \mid s_n^E(T^k x) = s\} \leq \left[\frac{n}{p_n(s)}\right] + 1.$$

As usual, here we define $[x] := \max\{n \in \mathbb{N} \mid n \leq x\}$. We will now see that Lemma 1.5 enables us to understand how the measure of $A_n^E(s)$ depends on the period of s.

LEMMA 1.6. *Let $\tau = (M, TM, \ldots, T^{n-1}M)$ be a Rokhlin tower, $E \in \mathcal{B}$ and $s \in \{0, 1\}^n$. Then*

$$\mu(A_n^E(s) \cap \tau) \leq \frac{2}{p_n(s)}\mu(\tau).$$

Proof. For $J \subset \{0, \ldots, n-1\}$ we define

$$M(J) := \{x \in M \mid \{k \in \{0, \ldots, n-1\} \mid s_n^E(T^k x) = s\} = J\}.$$

If $I, J \subset \{0, \ldots, n-1\}$ and $I \neq J$, then $M(J) \cap M(I) = \emptyset$. Furthermore,

$$A_n^E(s) \cap \tau = \bigcup_{J \subset \{0,\ldots,n-1\}} \bigcup_{k \in J} T^k(M(J)).$$

From this equality and Lemma 1.5, it follows easily that

$$\mu(A_n^E(s) \cap \tau) \leq \left(\left[\frac{n}{p_n(s)}\right] + 1\right) \sum_{J \subset \{0,\ldots,n-1\}} \mu(M(J)) \leq \frac{2}{p_n(s)}\mu(\tau).$$

□

Lemma 1.6 combined with Lemma 1.8 below shows how the concept of periodicity for 01-names can be used to obtain lower estimates for the entropy of the refined partitions $\alpha(E)_0^{n-1}$. First we need to mention a simple property of the entropy function f.

Remark 1.7. If $x_1, \ldots, x_n, \lambda > 0$ with $x_i \leq \lambda$, then

$$\sum_{i=1}^n f(x_i) \geq \frac{f(\lambda)}{\lambda} \sum_{i=1}^n x_i,$$

because the concavity of f implies that the function $x \mapsto f(x)/x$ is monotone decreasing on $(0, \infty)$.

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

The proof of the following lemma is an immediate consequence of Remark 1.7.

LEMMA 1.8. *Let β be a finite partition of X, $F \in \mathcal{B}$ and $\lambda > 0$, such that*

$$\mu(B \cap F) \leq \lambda \quad \text{for all } B \in \beta.$$

Then

$$H_F(\beta) \geq \mu(F) \log_2 \frac{1}{\lambda}.$$

We will use Lemma 1.8 typically with $\beta = \alpha(E)_0^{n-1}$, $F = \tau$ for some Rokhlin tower τ and $\lambda = 2\,\mu(\tau)/p$, where p will be a lower bound for the periods of 01-names in a sufficiently large subset of τ (see Lemma 1.6).

Often it will be useful to manipulate a given Rokhlin tower by 'cutting it down' or to consider only a certain subtower. The following definition introduces the necessary notation.

Definition 1.9. Let $\tau = (M, TM, \ldots, T^{n-1}M)$ be a Rokhlin tower of length n.

(a) If $p \in \mathbb{N}$ such that $p|n$, then we define

$$\frac{1}{p} * M := \bigcup_{k=0}^{p-1} T^{kn/p} M$$

$$\frac{1}{p} * \tau := \left(\frac{1}{p} * M, T\left(\frac{1}{p} * M\right), \ldots, T^{(n/p)-1}\left(\frac{1}{p} * M\right)\right).$$

(b) For $0 \leq i \leq k \leq n-1$ we define

$$\tau_i^k := (T^i M, \ldots, T^k M).$$

To complete our discussion in §1, we need to establish a few more purely technical results. In Lemmas 1.10 and 1.11 we will examine properties of subnames of 01-names. If $s = (s_0, \ldots, s_{n-1}) \in \{0,1\}^n$ and $0 \leq i \leq k \leq n-1$, then we call $s_i^k := (s_i, \ldots, s_k) \in \{0,1\}^{k-i+1}$ a *subname* of s. The following two lemmas are very simple consequences of the definition of periodicity for 01-names as given in Definition 1.4.

LEMMA 1.10. *If $s = (s_0, \ldots, s_{n-1}) \in \{0,1\}^n$ and $0 \leq i \leq k \leq n-1$, then*

$$p_{k-i+1}(s_i^k) \leq p_n(s).$$

LEMMA 1.11. *If $s = (s_0, \ldots, s_{n-1}) \in \{0,1\}^n$, $0 \leq i \leq k \leq n-1$ and $p_n(s) < (k-i+1)/2$, then*

$$p_n(s) = p_{k-i+1}(s_i^k).$$

Definition 1.12. Assume that $\tau = (M, TM, \ldots, T^{n-1}M)$ is a Rokhlin tower, $1 \leq p \leq q \leq n$ and $E \in \mathcal{B}$. Then we define

$$P_p^q(n,\ E) := \{x \in X \mid p \leq p_n(s_n^E(x)) \leq q\}$$

$$B_p^q(\tau,\ E) := P_p^q(n,\ E) \cap M.$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

This definition says that $P_p^q(n,\ E)$ is the set of all $x \in X$, which have 01-names of length n with periods ranging between p and q.

It will often be convenient to have 01-names of a certain period in the bottom level of a Rokhlin tower, and this is the motivation for the definition of $B_p^q(\tau,\ E)$. The following lemma explains how we can 'cut down' a tower τ of length $2n$ to the tower $\frac{1}{2} * \tau$ in order to obtain a lower bound for the measure of $B_p^q(\frac{1}{2} * \tau,\ E)$ depending on the measure of $P_p^q(2n,\ E) \cap \tau$.

LEMMA 1.13. *Let $\tau = (M, TM, \ldots, T^{2n-1}M)$ be a Rokhlin tower of length $2n$ and let $E \in \mathcal{B}$. If $0 \leq p \leq q < n/2$, then*

$$\frac{\mu(B_p^q(\frac{1}{2} * \tau,\ E))}{2\mu(M)} \geq \frac{\mu(P_p^q(2n,\ E) \cap \tau)}{\mu(\tau)} - \frac{1}{2}.$$

2. *Negative results for aperiodic transformations*

Our first objective in this section is to show that the entropy convergence rates of type $(LS < \infty)$, $(LS \leq c)$, $(LI < \infty)$ or $(LI \leq c)$ cannot be used to distinguish aperiodic systems with entropy zero if we choose $P(X)$ as the class of partitions. The assumption of aperiodicity will allow us to use the well-known Rokhlin lemma to construct Rokhlin towers of arbitrary size and length.

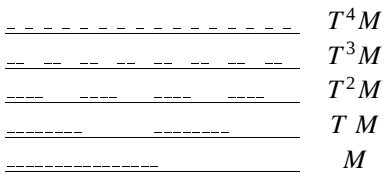
Theorem 2.8 will show that there are no aperiodic systems of type $(LI < \infty)$ for $(\{a_n\}_{n=1}^{\infty}, P(X))$ if $\lim_{n\to\infty} a_n/n = 0$. This implies, obviously, the non-existence of aperiodic systems of type $(LS < \infty)$, $(LS \leq c)$ or $(LI \leq c)$ under the same condition on $\{a_n\}_{n=1}^{\infty}$. The proof involves certain 'stability properties' of 'independent sets' in Rokhlin towers (see Definition 2.1 and Lemma 2.6).

Definition 2.1. Let $\tau = (M, TM, \ldots, T^{n-1}M)$ be a Rokhlin tower. Then we call $I \in \mathcal{B}$ *independent in* τ if

$$\{\{T^{-k}(I \cap T^k M), M \setminus T^{-k}(I \cap T^k M)\}\}_{k=0}^{n-1}$$

is a collection of independent partitions of M. For convenience we will always assume that $\mu(I \cap T^k M) = \mu(M)/2$ for all $k \in \{0, \ldots, n-1\}$.

Here is an example of an independent set in a Rokhlin tower of length 5, where the set $\tau \cap I$ is marked with '$\underline{\ \ }$':



Remark 2.2. If (X, T) is an aperiodic measure-preserving system, then $(X, \mathcal{B}, \mu)$ is obviously non-atomic and there are independent sets in any given Rokhlin tower $\tau \subset X$ (by a few measure-theoretic arguments).

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

Now we need some simple facts about binomial coefficients. For $n \in \mathbb{N}$ and $k \in \{0, \dots, n\}$ we define

$$a_n(k) := \sum_{i=0}^{k} \binom{n}{i} \quad \text{and} \quad b_n(k) := \sum_{i=0}^{k-1} i \binom{n}{i}.$$

As an easy consequence of these definitions we obtain the following lemma.

LEMMA 2.3. *If $n \geq 3$ and $k \in \{2, \dots, n-1\}$, then*

$$\frac{a_n(k+1)}{b_n(k+1)} \leq \frac{a_n(k)}{b_n(k)}.$$

LEMMA 2.4. *If $n \geq 16$ and $k \in \{[n/4]+1, \dots, n\}$, then*

$$\frac{a_n(k)}{b_n(k)} < \frac{256}{n}.$$

Proof. According to Lemma 2.3 it is sufficient to show the inequality for $k_0 := [n/4]+1$:

$$\frac{a_n(k_0)}{b_n(k_0)} \leq \frac{2\sum_{i=[k_0/2]}^{k_0} \binom{n}{i}}{[k_0/2]\sum_{i=[k_0/2]}^{k_0-1} \binom{n}{i}} \leq \frac{2}{[k_0/2]} \left(1 + \frac{\binom{n}{k_0}}{\binom{n}{k_0-1}}\right) = \frac{2(n+1)}{k_0[k_0/2]} < \frac{256}{n}.$$

□

Also very important for our purposes is the following inequality, which is a consequence of Stirling's estimate (see [**8**, p. 75]). If $\alpha \in (0,1)$ and $n \in \mathbb{N}$, then

$$a_n([\alpha n]) \leq \frac{\sqrt{n}}{(\alpha^\alpha (1-\alpha)^{1-\alpha})^n}.$$

Using this result, we obtain for all $k \in \{0, \dots, [n/4]\}$ that

$$\frac{a_n(k)}{2^n} \leq \sqrt{n}\, c^n \quad \text{where} \quad c := \frac{1}{2(1/4)^{1/4}(1-1/4)^{1-1/4}} < 1. \tag{2.5}$$

Now we will use Lemma 2.4 and 2.5 to prove a stability property for independent sets in a Rokhlin tower, which can be described as follows.

If $\tau = (M, TM, \dots, T^{n-1}M)$ is a Rokhlin tower and $I \subset \tau$ an independent set in τ, then $\{A_n^I(s) \cap M \mid s \in \{0,1\}^n\}$ is a partition of M into 2^n sets of equal measure $2^{-n}\mu(M)$. The question is how the measure of the atoms in this partition changes if I is replaced by a set $E \subset \tau$ such that $\mu(E \triangle I)$ is small relative to the measure of τ. In order to approach this question, we can think of the partition $\{A_n^E(s) \cap M \mid s \in \{0,1\}^n\}$ generated by E as being split up into 'good' and 'bad' atoms, where the 'good' ones are those which have not increased in measure by too large an amount compared to $2^{-n}\mu(M)$. More precisely, we say that an atom is 'good' if $\mu(A_n^E(s) \cap M) \leq \sqrt{n}\, c^n \mu(M)$. Lemma 2.6 shows that the 'bad' atoms cover less than half of M, or, positively speaking, the 'good' atoms cover more than half of M. This substantial supply of 'good' atoms will be used in the conclusion of the proof for Theorem 2.8 to generate refined partitions $\alpha(E)_0^{n-1}$ with large entropy values.

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

LEMMA 2.6. *Let $\tau = (M, TM, \dots, T^{n-1}M)$ be a Rokhlin tower with $n \geq 16$ and let $I \subset \tau$ be an independent set in τ. If $E \in \mathcal{B}$ is a subset of τ with $\mu(I \triangle E) < \mu(\tau)/512$,*

$$S := \{s \in \{0,1\}^n \mid \mu(A_n^E(s) \cap M) > \sqrt{n}\, c^n \mu(M)\} \quad \textit{and} \quad N := \bigcup_{s \in S} A_n^E(s) \cap M,$$

then

$$\mu(N) < \tfrac{1}{2}\mu(M).$$

Proof. For $s, r \in \{0,1\}^n$ we define

$$M(s,r) := A_n^E(s) \cap A_n^I(r) \cap M,$$

$$\tau(s,r) := \bigcup_{k=0}^{n-1} T^k(M(s,r)) \quad \text{and} \quad \tau(s) := \bigcup_{r \in \{0,1\}^n} \tau(s,r).$$

Then

$$\begin{aligned} \mu((I \triangle E) \cap \tau(s)) &= \sum_{r \in \{0,1\}^n} \mu((I \triangle E) \cap \tau(s,r)) \\ &= \sum_{r \in \{0,1\}^n} \|s - r\|_1 \mu(M(s,r)) = \sum_{i=0}^{n} \sum_{\substack{r \in \{0,1\}^n \\ \|s-r\|_1 = i}} i\,\mu(M(s,r)). \quad (1) \end{aligned}$$

Let

$$x_i(s) := \sum_{\substack{r \in \{0,1\}^n \\ \|s-r\|_1 = i}} \mu(M(s,r)) \quad \text{and} \quad y_i := \frac{1}{2^n} \binom{n}{i} \mu(M).$$

Since $M(s,r) \subset A_n^I(r) \cap M$ and $\mu(A_n^I(r) \cap M) = 2^{-n}\mu(M)$, we obtain

$$x_i(s) \leq \sum_{\substack{r \in \{0,1\}^n \\ \|s-r\|_1 = i}} \frac{\mu(M)}{2^n} = y_i. \quad (2)$$

Furthermore, since $\sum_{i=0}^{n} x_i(s) \leq \mu(A_n^E(s) \cap M) \leq \mu(M)$, we can find a $k_s \in \{0, \dots, n\}$ such that

$$\sum_{i=0}^{k_s} y_i = \frac{1}{2^n} \sum_{i=0}^{k_s} \binom{n}{i} \mu(M) \geq \sum_{i=0}^{n} x_i(s) \geq \frac{1}{2^n} \sum_{i=0}^{k_s - 1} \binom{n}{i} \mu(M) = \sum_{i=0}^{k_s - 1} y_i. \quad (3)$$

As a rather trivial consequence of (2) and (3) we have

$$\sum_{i=0}^{n} i\, x_i(s) \geq \sum_{i=0}^{k_s - 1} i\, y_i. \quad (4)$$

We will now show that $k_s > [n/4]$ for all $s \in S$. If $s \in S$, then $\mu(A_n^E(s) \cap M) > \sqrt{n}\, c^n \mu(M)$ and, therefore, (3) implies that

$$\frac{a_n(k_s)}{2^n} = \frac{1}{2^n} \sum_{i=0}^{k_s} \binom{n}{i} \geq \frac{\sum_{i=0}^{k_s} x_i(s)}{\mu(M)} \geq \frac{\mu(A_n^E(s) \cap M)}{\mu(M)} > \sqrt{n}\, c^n.$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

Since $2^{-n}a_n(k) \le \sqrt{n}\,c^n$ for all $k \in \{0, \dots, [n/4]\}$ (by 2.5), it follows that $k_s > [n/4]$. Now we obtain for all $s \in S$ that

$$\begin{aligned}\mu((I \triangle E) \cap \tau(s)) &= \sum_{i=0}^{n} i x_i(s) \quad \text{(by (1))} \quad \ge \sum_{i=0}^{k_s-1} i y_i \quad \text{(by (4))} \\ &= \frac{1}{2^n} b_n(k_s)\mu(M) > \frac{1}{2^n}\frac{n a_n(k_s)}{256}\mu(M) \quad \text{(by 2.4)} \\ &= \frac{n}{256}\frac{1}{2^n}\sum_{i=0}^{k_s}\binom{n}{i}\mu(M) \ge \frac{n}{256}\sum_{i=0}^{n} x_i(s) \quad \text{(by (3))} \qquad (5)\end{aligned}$$

Hence,

$$\begin{aligned}\mu(I \triangle E) &\ge \sum_{s\in S}\mu((I\triangle E)\cap\tau(s)) \ge \frac{n}{256}\sum_{s\in S}\sum_{i=0}^{n} x_i(s) \quad \text{(by (5))} \\ &= \frac{n}{256}\sum_{s\in S}\sum_{r\in\{0,1\}^n}\mu(M(s,r)) = \frac{n}{256}\mu(N).\end{aligned}$$

This implies that

$$\mu(N) \le \frac{256}{n}\mu(I\triangle E) < \frac{1}{2n}\mu(\tau) = \frac{1}{2}\mu(M),$$

because $\mu(I\triangle E) < \mu(\tau)/512$ by assumption. □

The following lemma shows how we can use the stability property of independent sets in order to obtain a lower estimate for the entropy of $\alpha(E)_0^{n-1}$, if E is a set that differs from a given independent set only by a small amount in measure.

We will use the well-known fact that if α and β are finite partitions of X, then $H(\alpha \vee \beta) \le H(\alpha) + H(\beta)$.

LEMMA 2.7. *Let $\tau = (M, TM, \dots, T^{2n-1}M)$ be a Rokhlin tower of length $2n$ with $n \ge 16$ and assume that $I \subset \tau$ is independent in τ. If $E \in \mathcal{B}$ is a subset of τ such that $\mu(I\triangle E) < \mu(\tau)/1024$, then*

$$H_{\tau_0^{n-1}}(\alpha(E)_0^{n-1}) \ge -\frac{\mu(\tau)}{4}\left(n\log_2 c + \frac{3}{2}\log_2 n - \log_2\frac{\mu(\tau)}{2}\right) - 1.$$

Proof. We define three partitions of X as follows:

$$\begin{aligned}\alpha &:= \{T^k M\}_{k=0}^{n-1} \cup \{X \setminus \tau_0^{n-1}\}, \\ \beta &:= \{T^k M \cap A_n^E(s) \mid k \in \{0,\dots,n-1\} \quad \text{and} \quad s \in \{0,1\}^n\} \cup \{X \setminus \tau_0^{n-1}\} \quad \text{and} \\ \gamma &:= \{A_n^E(s) \cap \tau_0^{n-1} \mid s \in \{0,1\}^n\} \cup \{X \setminus \tau_0^{n-1}\}.\end{aligned}$$

Then we have $\beta = \alpha \vee \gamma$ and, therefore,

$$\begin{aligned}H_{\tau_0^{n-1}}(\alpha(E)_0^{n-1}) &= H(\gamma) - f(\mu(X \setminus \tau_0^{n-1})) \\ &\ge H(\beta) - H(\alpha) - f(\mu(X \setminus \tau_0^{n-1})) \\ &\ge H(\beta) - nf(\mu(M)) - 2f(\mu(X \setminus \tau_0^{n-1})). \qquad (1)\end{aligned}$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

For $k \in \{0, \dots, n-1\}$ we define

$$S_k := \{s \in \{0,1\}^n \mid \mu(T^k M \cap A_n^E(s)) > \sqrt{n}\, c^n \mu(M)\} \quad \text{and}$$
$$N_k := \bigcup_{s \in S_k} T^k M \cap A_n^E(s).$$

Then

$$\mu((I \cap \tau_k^{k+n-1}) \triangle (E \cap \tau_k^{k+n-1})) \le \mu(I \triangle E) < \frac{\mu(\tau)}{1024} = \frac{\mu(\tau_k^{k+n-1})}{512}$$

and, therefore, we can apply Lemma 2.6 to $(I \cap \tau_k^{k+n-1})$, $(E \cap \tau_k^{k+n-1})$ and τ_k^{k+n-1} to conclude that

$$\mu(N_k) < \tfrac{1}{2}\mu(M) \quad \text{for all } k \in \{0, \dots, n-1\}. \tag{2}$$

Hence,

$$\begin{aligned}
H(\beta) &\ge \sum_{k=0}^{n-1} \sum_{s \in \{0,1\}^n \setminus S_k} f(\mu(T^k M \cap A_n^E(s))) + f(\mu(X \setminus \tau_0^{n-1})) \\
&\ge \sum_{k=0}^{n-1} \frac{f(\sqrt{n}\, c^n \mu(M))}{\sqrt{n}\, c^n \mu(M)} \sum_{s \in \{0,1\}^n \setminus S_k} \mu(T^k M \cap A_n^E(s)) + f(\mu(X \setminus \tau_0^{n-1})) \quad \text{(by 1.7)} \\
&= -\log_2(\sqrt{n}\, c^n \mu(M)) \sum_{k=0}^{n-1} (\mu(M) - \mu(N_k)) + f(\mu(X \setminus \tau_0^{n-1})) \\
&> -\log_2(\sqrt{n}\, c^n \mu(M)) \frac{n}{2} \mu(M) + f(\mu(X \setminus \tau_0^{n-1})) \quad \text{(by (2))} \\
&= -\frac{\mu(\tau)}{4} \left(n \log_2 c + \frac{1}{2} \log_2 n - \log_2 \mu(M) \right) + f(\mu(X \setminus \tau_0^{n-1})). \qquad (3)
\end{aligned}$$

By using (1) and (3) we obtain

$$\begin{aligned}
H_{\tau_0^{n-1}}(\alpha(E)_0^{n-1}) &\ge -\frac{\mu(\tau)}{4} \left(n \log_2 c + \frac{1}{2} \log_2 n - \log_2 \mu(M) \right) \\
&\quad -nf(\mu(M)) - f(\mu(X \setminus \tau_0^{n-1})) \\
&\ge -\frac{\mu(\tau)}{4} \left(n \log_2 c + \frac{1}{2} \log_2 n - \log_2 \mu(M) \right) \\
&\quad + \frac{\mu(\tau)}{2} \log_2 \mu(M) - 1 \\
&= -\frac{\mu(\tau)}{4} \left(n \log_2 c + \frac{3}{2} \log_2 n - \log_2 \frac{\mu(\tau)}{2} \right) - 1.
\end{aligned}$$

□

THEOREM 2.8. *If (X, T) is an aperiodic measure-preserving system and $\{a_n\}_{n=1}^{\infty}$ is a positive monotone increasing sequence with $\lim_{n\to\infty} a_n/n = 0$, then there exists an $E \in \mathcal{B}$ such that*

$$\lim_{n\to\infty} \frac{H(\alpha(E)_0^{n-1})}{a_n} = \infty.$$

In particular, (X, T) is not of type $(LI < \infty)$ for $(\{a_n\}_{n=1}^{\infty}, P(X))$.

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

Proof. For $n \in \mathbb{N}$, we define $\delta_n := 4096^{-n}$ and choose an $N_n \in \mathbb{N}$ such that

$$\frac{a_m}{m} < \frac{\delta_n}{n} \quad \text{for all } m \geq N_n. \tag{1}$$

W.l.o.g. we may assume that

$$N_n \geq 16, \tag{2}$$

$$\lim_{n\to\infty} \frac{n}{N_n} = 0 \quad \text{and} \tag{3}$$

$$N_{n+1} > N_n \tag{4}$$

for all $n \in \mathbb{N}$.

We define recursively a sequence of sets $E_n \in \mathcal{B}$ as follows. Let $E_0 := \emptyset$ and assume that E_{n-1} has been defined. We choose a Rokhlin tower

$$\tau_n = (M_n, TM_n, \ldots, T^{2N_{n+1}-1}M_n)$$

of length $2N_{n+1}$ with $\mu(\tau_n) = \delta_n$. Let $I_n \subset \tau_n$ be independent in τ_n and let $E_n := (E_{n-1} \setminus \tau_n) \cup I_n$. Then $\mu(E_n \triangle E_{n-1}) \leq \mu(\tau_n) = \delta_n$ and, therefore, $\{\chi_{E_n}\}_{n=0}^{\infty}$ is a Cauchy sequence in $L_1(X)$ converging to χ_E for some $E \in \mathcal{B}$.

Now let $i \in \{N_n, \ldots, N_{n+1} - 1\}$ (see (4)) and let us choose a $K_i \in \mathbb{N}$ such that $i\,K_i \leq N_{n+1} < i(K_i + 1)$. We define

$$M_n(i) := \bigcup_{k=0}^{K_i-1} T^{2ik} M_n \quad \text{and} \quad \tau_n(i) := (M_n(i), TM_n(i), \ldots, T^{2i-1}M_n(i)).$$

Then, $\tau_n(i)$ is a Rokhlin tower of length $2i$ with $\mu(\tau_n(i)) \geq \frac{1}{2}\mu(\tau_n)$ and I_n is independent in $\tau_n(i)$. Furthermore,

$$\mu((I_n \triangle E) \cap \tau_n(i)) \leq \sum_{k=n+1}^{\infty} \mu(\tau_k) < 2\delta_{n+1} = \frac{\mu(\tau_n)}{2048} = \frac{\mu(\tau_n(i))}{1024}.$$

Using this estimate and (2), we can apply Lemma 2.7 and obtain

$$H_{\tau_n(i)_0^{i-1}}(\alpha(E)_0^{i-1}) \geq -\frac{\mu(\tau_n(i)_0^{i-1})}{4}\left(i\log_2 c + \frac{3}{2}\log_2 i - \log_2 \frac{\mu(\tau_n(i)_0^{i-1})}{2}\right) - 1. \tag{5}$$

Now we define n_i to be the unique integer with $N_{n_i} \leq i < N_{n_i+1}$ (w.l.o.g. $i \geq N_1$). Then

$$\lim_{i\to\infty} \frac{n_i}{i} = 0 \quad \text{(by (3))} \tag{6}$$

and

$$\begin{aligned}
H(\alpha(E)_0^{i-1}) &\geq H_{\tau_n(i)_0^{i-1}}(\alpha(E)_0^{i-1}) - 2 \quad \text{(by 1.2)} \\
&\geq -\frac{\mu(\tau_n(i)_0^{i-1})}{4}\left(i\log_2 c + \frac{3}{2}\log_2 i - \log_2 \frac{\mu(\tau_n(i)_0^{i-1})}{2}\right) - 3 \quad \text{(by (5))} \\
&\geq -i\,\frac{\delta_{n_i}}{8}\log_2 c - \frac{\delta_{n_i}}{4}\left(\frac{3}{2}\log_2 i - \log_2 \frac{\delta_{n_i}}{4}\right) - 3 \\
&= -i\,\frac{\delta_{n_i}}{8}\left(\log_2 c + 3\frac{\log_2 i}{i} + \frac{4}{i} + 24\frac{n_i}{i}\right) - 3.
\end{aligned}$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

This implies by using (6) that for sufficiently large i we have

$$H(\alpha(E)_0^{i-1}) \geq -i\frac{\delta_{n_i}}{16}\log_2 c - 3.$$

Hence,

$$\begin{aligned}\lim_{i\to\infty}\frac{H(\alpha(E)_0^{i-1})}{a_i} &\geq \lim_{i\to\infty}\left(-\frac{i}{a_i}\delta_{n_i}\frac{\log_2 c}{16} - \frac{3}{a_i}\right)\\ &\geq \lim_{i\to\infty}\left(-n_i\frac{\log_2 c}{16} - \frac{3}{a_i}\right) \quad \text{(by (1))}\\ &= \infty\end{aligned}$$

□

3. *Positive results for aperiodic transformations*

Our first objective in this section is to prove a very general positive result about the convergence type ($LS = \infty$) for aperiodic systems. This will be the only time that we do not choose $P(X)$ as the class of partitions, but instead the following variation of it:

$$R(X) := \left\{\alpha(E) \mid E \in \mathcal{B} \quad \text{and} \quad \lim_{n\to\infty}\max\{\mu(A) \mid A \in \alpha(E)_0^{n-1}\} = 0\right\}.$$

In order to clarify our assumptions, we want to mention the following fact.

Remark 3.1. If (X, T) is a measure-preserving system such that $R(X) \neq \emptyset$, then (X, T) is aperiodic. This follows with a few simple arguments from the observation that every periodic point x of T generates 01-names with periods smaller than or equal to the period of x under T.

It is trivial to see that $R(X)$ is still 'large enough' in the sense that any convergence type defined with $R(X)$ is an isomorphism invariant. The reason for choosing $R(X)$ is that in some aperiodic systems we can find sets $E \in \mathcal{B}$, such that $\alpha(E)$ is a non-trivial partition and for some $N \in \mathbb{N}$ we have $\alpha(E)_0^{n-1} = \alpha(E)_0^{N-1}$ for all $n \geq N$, i.e. the $\alpha(E)_0^{n-1}$ are stable for large enough n. This is, for example, the case for the von Neumann–Kakutani adding machine defined on $[0, 1]$, where $E := [0, 1/2]$ is such a set. The way in which we will use the condition

$$\lim_{n\to\infty}\max\{\mu(A) \mid A \in \alpha(E)_0^{n-1}\} = 0$$

is expressed in the following lemma, for which we will need the notation of Definition 1.12.

LEMMA 3.2. *If (X, T) is an aperiodic measure-preserving system and $E \in \mathcal{B}$ such that*

$$\lim_{n\to\infty}\max\{\mu(A) \mid A \in \alpha(E)_0^{n-1}\} = 0,$$

then

$$\lim_{n\to\infty}\mu(P_1^q(n,\ E)) = 0 \quad \textit{for all } q \in \mathbb{N}.$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

Proof. The definition of periodicity for 01-names as given in Definition 1.4 implies trivially that every $s \in \{0,1\}^q$ is completely determined by its first $p_q(s)$ coordinates. This means that $P_1^q(n,\ E)$ is the union of at most 2^q atoms of $\alpha(E)_0^{n-1}$ for all $n \in \mathbb{N}$. Since q is constant, the condition $\lim_{n\to\infty} \max\{\mu(A) \mid A \in \alpha(E)_0^{n-1}\} = 0$ implies immediately that $\lim_{n\to\infty} \mu(P_1^q(n,\ E)) = 0$. □

In Theorem 3.9 we will show that every aperiodic system is of type $(LS = \infty)$ for $(\{g(\log_2 n)\}_{n=1}^{\infty},\ R(X))$ whenever $g : [0,\infty) \to \mathbb{R}$ is a positive monotone increasing function with $\int_1^{\infty} g(x)/x^2\,dx < \infty$. As a corollary of this theorem we will see that every completely ergodic system (X, T) is of type $(LS = \infty)$ for $(\{g(\log_2 n)\}_{n=1}^{\infty},\ P(X))$ under the same conditions on g. (We say that T is completely ergodic if T^n is ergodic for all $n \in \mathbb{Z}$.)

Surprisingly, the proof of Theorem 3.9 is based on a rather simple analytic observation which is stated in Lemma 3.4. For the proof of Lemma 3.4 we will need the following fact, which is very easy to prove.

LEMMA 3.3. *If* $\{a_n\}_{n=1}^{\infty}$ *is a positive monotone increasing sequence such that* $\sum_{n=1}^{\infty} a_n/(n(n+1)) < \infty$, *then*

$$\lim_{n\to\infty} \frac{a_n}{n} = 0.$$

LEMMA 3.4. *Let* $\{a_n\}_{n=1}^{\infty}$ *be a positive monotone increasing sequence such that*

$$d := \sum_{n=1}^{\infty} \frac{a_n}{n(n+1)} < \infty.$$

For $n \in \mathbb{N}$, $\lambda > 0$, $m \in \{1,\dots,n\}$ *and* $x = (x_1,\dots,x_n) \in \mathbb{R}^n$ *we define*

$$\Delta_n(\lambda) := \left\{ x = (x_1,\dots,x_n) \in \mathbb{R}^n \mid x_k \geq 0 \quad \textit{and} \quad \sum_{k=1}^{n} x_k \geq \lambda \right\} \quad \textit{and}$$

$$s_m^n(x) := \frac{1}{a_m} \sum_{k=1}^{m} k\,x_k.$$

Then

$$\liminf_{n\to\infty} \inf_{x\in\Delta_n(\lambda)} \max_{m\in\{1,\dots,n\}} s_m^n(x) \geq \frac{\lambda}{d}.$$

Proof. Let us assume that $\liminf_{n\to\infty} \inf_{x\in\Delta_n(\lambda)} \max_{m\in\{1,\dots,n\}} s_m^n(x) < \lambda/d$. Then there exists an $\varepsilon > 0$, a strictly increasing sequence $\{n_i\}_{i=1}^{\infty} \subset \mathbb{N}$ and a sequence $\{x(i)\}_{i=1}^{\infty}$ such that

$$d\varepsilon < 1, \tag{1}$$

$$x(i) \in \Delta_{n_i}(\lambda) \quad \text{and} \tag{2}$$

$$s_m^{n_i}(x(i)) < \lambda\left(\frac{1}{d} - \varepsilon\right) =: c_\varepsilon \tag{3}$$

for all $i \in \mathbb{N}$ and $m \in \{1,\dots,n_i\}$. Then we obtain

$$\sum_{m=1}^{n_i} \frac{a_m}{m(m+1)} \quad > \quad \frac{1}{c_\varepsilon} \sum_{m=1}^{n_i} \frac{\sum_{k=1}^{m} k x(i)_k}{m(m+1)} \quad \text{(by (3))}$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

$$= \frac{1}{c_\varepsilon}\sum_{k=1}^{n_i} k\,x(i)_k \sum_{m=k}^{n_i}\frac{1}{m(m+1)} = \frac{1}{c_\varepsilon}\sum_{k=1}^{n_i} k\,x(i)_k\left(\frac{1}{k}-\frac{1}{n_i+1}\right)$$

$$\geq \frac{1}{c_\varepsilon}\left(\lambda - \frac{1}{n_i+1}\sum_{k=1}^{n_i} k x(i)_k\right) \quad \text{(by (2))} > \frac{\lambda}{c_\varepsilon} - \frac{a_{n_i}}{n_i+1} \quad \text{(by (3))}.$$

Using Lemma 3.3 and (1), this estimate implies that

$$d = \lim_{i\to\infty}\sum_{m=1}^{n_i}\frac{a_m}{m(m+1)} \geq \lim_{i\to\infty}\left(\frac{\lambda}{c_\varepsilon} - \frac{a_{n_i}}{n_i+1}\right) = \frac{\lambda}{c_\varepsilon} = \frac{d}{1-d\varepsilon} > d$$

which is a contradiction, and the proof is complete. □

Before we can see how Lemma 3.4 can be useful for our purposes, we need to set up some more notation.

Definition 3.5. Let $E \in \mathcal{B}$ and let $\tau = (M, TM, \ldots, T^{2^n-1}M)$ be a Rokhlin tower of length 2^n. For $k \in \{1, \ldots, n-1\}$ we define subtowers of τ of length 2^{n-1} as follows (see Definition 1.12):

$$\tau_k(\tau, E) := (B_{2^{k-1}}^{2^k-1}(\tau,\ E), TB_{2^{k-1}}^{2^k-1}(\tau,\ E), \ldots, T^{2^{n-1}-1}B_{2^{k-1}}^{2^k-1}(\tau,\ E)).$$

This definition says that $\tau_k(\tau, E)$ is a subtower of τ beginning at the bottom level of τ, which is contained in the lower half of τ, i.e. in $\tau_0^{2^{n-1}-1}$. Furthermore, all points in the bottom level of $\tau_k(\tau, E)$ have 01-names of length 2^n with periods ranging between 2^{k-1} and $2^k - 1$.

Remark 3.6. Let $E \in \mathcal{B}$ and let $\tau = (M, TM, \ldots, T^{2^n-1}M)$ be a Rokhlin tower of length 2^n. If $n \geq 3$, $1 \leq k < j < n$, $x \in B_{2^{k-1}}^{2^k-1}(\tau,\ E)$ and $i \in \{0, \ldots, 2^{n-1}-1\}$, then $s_{2^j}^E(T^i x)$ is a subname of $s_{2^n}^E(x)$ of length 2^j. Since $k < j$, we have $2^k - 1 < 2^j/2$ and, therefore, Lemma 1.11 implies that $p_{2^j}(s_{2^j}^E(T^i x)) = p_{2^n}(s_{2^n}^E(x))$. Hence,

$$2^{k-1} \leq p_{2^j}(s_{2^j}^E(T^i x)) \leq 2^k - 1 \quad \text{for all } x \in \tau_k(\tau, E).$$

Using Lemma 1.6, Lemma 1.8 and Remark 3.6 we will now find a lower estimate for the entropy of $\alpha(E)_0^{2^j-1}$ restricted to $\tau_k(\tau, E)$ if $1 \leq k < j < n$.

LEMMA 3.7. *Let $E \in \mathcal{B}$ and let $\tau = (M, TM, \ldots, T^{2^n-1}M)$ be a Rokhlin tower of length 2^n. If $n \geq 3$ and $1 \leq k < j < n$, then*

$$H_{\tau_k(\tau,E)}(\alpha(E)_0^{2^j-1}) \geq (k-2)\mu(\tau_k(\tau, E)).$$

Proof. First we apply Lemma 1.6 to the tower $(1/2^{n-j-1}) * \tau_k(\tau, E)$ and obtain for all $s \in \{0,1\}^{2^j}$ that

$$\mu(A_{2^j}^E(s) \cap \tau_k(\tau, E)) = \mu\left(A_{2^j}^E(s) \cap \frac{1}{2^{n-j-1}} * \tau_k(\tau, E)\right)$$

$$\leq \frac{2}{p_{2^j}(s)}\mu\left(\frac{1}{2^{n-j-1}} * \tau_k(\tau, E)\right) = \frac{2}{p_{2^j}(s)}\mu(\tau_k(\tau, E)).$$

Therefore, Remark 3.6 implies that

$$\mu(A_{2^j}^E(s_{2^j}^E(x)) \cap \tau_k(\tau, E)) \leq \frac{1}{2^{k-2}}\mu(\tau_k(\tau, E)) \quad \text{for all } x \in \tau_k(\tau, E).$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

This means that $\alpha(E)_0^{2^j-1}$ induces a partition in $\tau_k(\tau, E)$ for which every atom has measure less than or equal to $2^{-(k-2)}\mu(\tau_k(\tau, E))$. Using Lemma 1.8 we can conclude that

$$H_{\tau_k(\tau,E)}(\alpha(E)_0^{2^j-1}) \geq \mu(\tau_k(\tau, E)) \log_2 \frac{2^{k-2}}{\mu(\tau_k(\tau, E))} \geq (k-2)\mu(\tau_k(\tau, E)).$$

□

The relation between Lemma 3.4 and Lemma 3.7 can be described as follows. In Lemma 3.4 we made a statement about sums of the form $\sum_{k=1}^n kx_k$ with $\sum_{k=1}^n x_k \geq \lambda$. In the proof of Proposition 3.8 we will construct a Rokhlin tower τ of length 2^{n+1} and define x_k to be $\mu(\tau_{k-1}(\frac{1}{2}*\tau, E))$. With this definition the terms $k\,x_k$ appear in Lemma 3.7 as part of a lower estimate for the entropy of $\alpha(E)_0^{2^j-1}$ restricted to $\tau_{k-1}(\frac{1}{2} * \tau, E)$. Therefore, $\sum_{k=1}^n k\,x_k$ will become a lower estimate for the entropy of $\alpha(E)_0^{2^j-1}$ restricted to the union of the towers $\tau_{k-1}(\frac{1}{2} * \tau, E)$. Furthermore, the constant λ will be interpreted as a lower bound for the total measure of the towers $\tau_{k-1}(\frac{1}{2} * \tau, E)$ and in order to find such a lower bound, we will use Lemma 1.13.

PROPOSITION 3.8. *Let (X, T) be an aperiodic measure-preserving system and let $\{a_n\}_{n=1}^\infty$ be a positive monotone increasing sequence such that*

$$d := \sum_{n=1}^{\infty} \frac{a_n}{n(n+1)} < \infty.$$

If $E \in \mathcal{B}$ and

$$\lim_{n\to\infty} \max\{\mu(A) \mid A \in \alpha(E)_0^{n-1}\} = 0,$$

then

$$\limsup_{n\to\infty} \frac{H(\alpha(E)_0^{2^n-1})}{a_n} > \frac{1}{128d}.$$

Proof. We may assume that $\lim_{n\to\infty} a_n = \infty$, otherwise the statement is trivial. Let $E \in \mathcal{B}$ such that

$$\lim_{n\to\infty} \max\{\mu(A) \mid A \in \alpha(E)_0^{n-1}\} = 0. \tag{1}$$

If $\lambda := 1/32$, then it is obviously sufficient to show that for every $N \in \mathbb{N}$ there exists a $j \in \mathbb{N}$ such that

$$j \geq N \quad \text{and} \quad \frac{H(\alpha(E)_0^{2^j-1})}{a_j} \geq \frac{\lambda}{4d}. \tag{2}$$

Using Lemma 3.3, the assumption $\lim_{n\to\infty} a_n = \infty$, Lemma 3.4 and (1), we may choose N so large that for all $n \geq N$ we have

$$\frac{1}{8}\frac{n-2}{a_{n+1}} > \frac{\lambda}{4d} \quad \text{and} \quad \frac{5}{a_n} < \frac{\lambda}{4d}, \tag{3}$$

$$\inf_{x\in\Delta_n(\lambda)} \max_{m\in\{1,\dots,n\}} s_m^n(x) > \frac{\lambda}{2d} \quad \text{and} \tag{4}$$

$$\max\{\mu(A) \mid A \in \alpha(E)_0^{n-1}\} < \frac{\lambda}{2^{2^N-1}}. \tag{5}$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

As we will see, inequality (5) will guarantee that the set of all $x \in X$ which have 01-names with periods smaller than 2^N is sufficiently small and this will allow us to choose $j \geq N$ as required in (2). Now let $n > N+1$ and let $\tau = (M, TM, \ldots, T^{2^{n+1}-1}M)$ be a Rokhlin tower of length 2^{n+1} such that

$$\mu(\tau) > \tfrac{1}{2}. \tag{6}$$

We consider two cases, depending on the measure of $P := P_{2^{n-1}}^{2^{n+1}}(2^{n+1},\ E) \cap \tau$.

Case 1. $\mu(P) \geq \mu(\tau)/4$. Using Lemma 1.6 we obtain that

$$\mu(A_{2^{n+1}}^E(s_{2^{n+1}}^E(x)) \cap \tau) \leq \frac{1}{2^{n-2}}\mu(\tau) \quad \text{for all } x \in P \tag{7}$$

and

$$\begin{aligned} H(\alpha(E)_0^{2^{n+1}}) &\geq H_P(\alpha(E)_0^{2^{n+1}-1}) - 2 \quad \text{(by 1.2)} \\ &\geq \mu(P)\log_2 \frac{2^{n-2}}{\mu(\tau)} \quad \text{(by 1.8 and (7))} \\ &\geq \tfrac{1}{4}\mu(\tau)(n-2) > \tfrac{1}{8}(n-2) \quad \text{(by (6)).} \end{aligned}$$

Therefore,

$$\frac{H(\alpha(E)_0^{2^{n+1}-1})}{a_{n+1}} > \frac{1}{8}\frac{n-2}{a_{n+1}} > \frac{\lambda}{4d} \quad \text{(by (3)).}$$

So in case 1 we see that (2) is satisfied for $j := n+1 > N$.

Case 2. $\mu(P) < \mu(\tau)/4$. In this case we have

$$\mu(P_1^{2^{n-1}-1}(2^{n+1},\ E) \cap \tau) = \mu(\tau \setminus P) > \tfrac{3}{4}\mu(\tau)$$

and, therefore, we obtain

$$\frac{\mu(B_1^{2^{n-1}-1}(\frac{1}{2} * \tau,\ E))}{2\mu(M)} \geq \frac{\mu(P_1^{2^{n-1}-1}(2^{n+1},\ E) \cap \tau)}{\mu(\tau)} - \frac{1}{2} > \frac{1}{4} \quad \text{(by 1.13).} \tag{8}$$

Since $B_1^{2^{n-1}-1}(\frac{1}{2} * \tau,\ E)$ is a disjoint union of the sets $B_1^1(\frac{1}{2} * \tau,\ E)$, $B_2^3(\frac{1}{2} * \tau,\ E)$, $B_4^7(\frac{1}{2} * \tau,\ E), \ldots, B_{2^{n-2}}^{2^{n-1}-1}(\frac{1}{2} * \tau,\ E)$, the towers $\tau_1(\frac{1}{2} * \tau, E), \ldots, \tau_{n-1}(\frac{1}{2} * \tau, E)$ are also pairwise disjoint. This shows that for $\tau_u := \bigcup_{k=1}^{n-1} \tau_k(\frac{1}{2} * \tau, E)$ we have

$$\begin{aligned} \mu(\tau_u) &= \sum_{k=1}^{n-1} \mu(\tau_k(\tfrac{1}{2} * \tau, E)) = 2^{n-1} \sum_{k=1}^{n-1} \mu(B_{2^{k-1}}^{2^k-1}(\tfrac{1}{2} * \tau,\ E)) \\ &= 2^{n-1}\mu(B_1^{2^{n-1}-1}(\tfrac{1}{2} * \tau,\ E)) > 2^{n-2}\mu(M) \quad \text{(by (8))} \\ &> \tfrac{1}{16} \quad \text{(by (6)).} \end{aligned} \tag{9}$$

Now let $x_1 := 0$ and for $k \in \{2, \ldots, n\}$ let

$$x_k := \mu(\tau_{k-1}(\tfrac{1}{2} * \tau, E)) \quad \text{and} \quad x := (0, \ldots, 0, x_{N+2}, \ldots, x_n) \in \mathbb{R}^n.$$

We wish to show that $x \in \Delta_n(\lambda)$. First we notice that

$$\#\{s \in \{0,1\}^{2^{n-1}} \mid p_{2^{n-1}}(s) \leq 2^N - 1\} = 2^{2^N - 1}, \tag{10}$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

because every $s \in \{0,1\}^{2^{n-1}}$ is uniquely determined by its first $p_{2^{n-1}}(s)$ coordinates (see Definition 1.4). Therefore,

$$\begin{aligned}\sum_{k=N+2}^{n} x_k &= \sum_{k=1}^{n} x_k - \sum_{k=1}^{N+1} x_k = \mu(\tau_u) - \sum_{k=1}^{N+1} \mu(\tau_{k-1}(\tfrac{1}{2} * \tau, E)) \\ &\geq \mu(\tau_u) - \sum_{\substack{s\in\{0,1\}^{2^{n-1}} \\ p_{2^{n-1}}(s)\leq 2^N-1}} \mu(A^E_{2^{n-1}}(s)) \quad \text{(by 3.6)} \\ &> \frac{1}{16} - 2^{2^N-1}\frac{\lambda}{2^{2^N-1}} \quad \text{(by (5), (9) and (10)).} \\ &= \lambda\end{aligned}$$

This shows that $x \in \Delta_n(\lambda)$.

According to (4), we can find a $j \in \{1, \ldots, n\}$ such that

$$s^n_j(x) > \frac{\lambda}{2d}. \tag{11}$$

Since the first $N+1$ coordinates of x are equal to zero, we have $s^n_m(x) = 0$ for all $m \in \{1, \ldots, N+1\}$ and this shows that $j \geq N+2$. So j satisfies the first condition in (2). With $\tau_j := \bigcup_{k=1}^{j-1} \tau_k(\frac{1}{2} * \tau, E)$ we obtain

$$\begin{aligned}H(\alpha(E)_0^{2^j-1}) &\geq H_{\tau_j}(\alpha(E)_0^{2^j-1}) - 2 \quad \text{(by 1.2)} \\ &= \sum_{s\in\{0,1\}^{2^j}} f(\mu(A^E_{2^j}(s) \cap \tau_j)) - 2 \\ &= \sum_{k=1}^{j-1} \sum_{\substack{s\in\{0,1\}^{2^j} \\ 2^{k-1}\leq p_{2^j}(s)\leq 2^k-1}} f(\mu(A^E_{2^j}(s) \cap \tau_j)) - 2 \quad \text{(by 3.6)} \\ &= \sum_{k=1}^{j-1} H_{\tau_k(\frac{1}{2}*\tau,E)}(\alpha(E)_0^{2^j-1}) - 2 \quad \text{(by 3.6)} \\ &\geq \sum_{k=1}^{j-1} (k-2)\mu(\tau_k(\tfrac{1}{2} * \tau, E)) - 2 \quad \text{(by 3.7)} \\ &= \sum_{k=1}^{j} (k-3)x_k - 2 \geq a_j s^n_j(x) - 3\sum_{k=1}^{j} x_k - 2 > a_j\frac{\lambda}{2d} - 5 \quad \text{(by (11)).}\end{aligned}$$

Finally, we obtain

$$\frac{H(\alpha(E)_0^{2^j-1})}{a_j} > \frac{\lambda}{2d} - \frac{5}{a_j} > \frac{\lambda}{4d} \quad \text{(by (3)).}$$

So, also in case 2 we have found a j satisfying (2) and the proof is complete. □

THEOREM 3.9. *Let (X, T) be an aperiodic measure-preserving system and let $g : [0, \infty) \to \mathbb{R}$ be a positive monotone increasing function with $\int_1^\infty g(x)/x^2\,dx < \infty$. If $E \in \mathcal{B}$ such that*

$$\lim_{n\to\infty} \max\{\mu(A) \mid A \in \alpha(E)_0^{n-1}\} = 0,$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

then

$$\limsup_{n\to\infty} \frac{H(\alpha(E)_0^{n-1})}{g(\log_2 n)} = \infty.$$

In other words, (X, T) *is of type* $(LS = \infty)$ *for* $(\{g(\log_2 n)\}_{n=1}^{\infty}, R(X))$.

Proof. If $a_n := g(n)$, then the condition $\int_1^{\infty} g(x)/x^2\,dx < \infty$ implies obviously that $\sum_{n=1}^{\infty} a_n/(n(n+1)) < \infty$. Using this fact, it is easy to find a positive monotone increasing sequence $\{b_n\}_{n=1}^{\infty}$ such that

$$\lim_{n\to\infty} \frac{b_n}{a_n} = \infty \quad \text{and} \quad d := \sum_{n=1}^{\infty} \frac{b_n}{n(n+1)} < \infty. \tag{1}$$

According to Proposition 3.8 we can find a strictly increasing sequence $\{m_j\}_{j=1}^{\infty} \subset \mathbb{N}$ such that

$$\frac{H(\alpha(E)_0^{2^{m_j}-1})}{b_{m_j}} > \frac{1}{256d} \quad \text{for all } j \in \mathbb{N}. \tag{2}$$

Hence,

$$\begin{aligned} \limsup_{n\to\infty} \frac{H(\alpha(E)_0^{n-1})}{g(\log_2 n)} &\geq \limsup_{j\to\infty} \frac{H(\alpha(E)_0^{2^{m_j}-1})}{a_{m_j}} \\ &= \limsup_{j\to\infty} \frac{H(\alpha(E)_0^{2^{m_j}-1})}{b_{m_j}} \frac{b_{m_j}}{a_{m_j}} \\ &= \infty \quad \text{(by (1) and (2)).} \end{aligned}$$

□

A proof for the following two corollaries can be found in [**1**].

COROLLARY 3.10. *Let* (X, T) *be ergodic and let* g *be given as in Theorem 3.9. If* $E \in \mathcal{B}$ *such that* $\alpha(E)_0^{n-1} \neq \alpha(E)_0^n$ *for all* $n \in \mathbb{N}$, *then*

$$\limsup_{n\to\infty} \frac{H(\alpha(E)_0^{n-1})}{g(\log_2 n)} = \infty.$$

(We say that two partitions are equal if they differ only on a set of measure zero. The condition $\alpha(E)_0^{n-1} \neq \alpha(E)_0^n$ *has to be interpreted in this way.)*

COROLLARY 3.11. *Let* (X, T) *be completely ergodic and let* g *be given as in* 3.9. *If* $E \in \mathcal{B}$ *such that* $0 < \mu(E) < 1$, *then*

$$\limsup_{n\to\infty} \frac{H(\alpha(E)_0^{n-1})}{g(\log_2 n)} = \infty.$$

In other words, every completely ergodic system is of type $(LS = \infty)$ *for* $(\{g(\log_2 n)\}_{n=1}^{\infty}, P(X))$.

We wish to make some comments on the results 3.9–3.11. One obvious question to ask is whether the condition $\int_1^{\infty} g(x)/x^2\,dx < \infty$ is the best possible one. In [**2**] we will

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

see that this is indeed the case. In fact, we will even show that if $\int_1^\infty g(x)/x^2\,dx = \infty$, then there exists a weakly mixing system (X, T) and a set $E \in \mathcal{B}$ such that $0 < \mu(E) < 1$ and $\lim_{n\to\infty} H(\alpha(E)_0^{n-1})/g(\log_2 n) = 0$. Notice that every weakly mixing system is, in particular, completely ergodic and, therefore, the conclusion of Corollary 3.11 is true for every weakly mixing system. Another question to ask is whether a statement analogous to Theorem 3.9 is possible for convergence rates of type $(LI = \infty)$ or at least $(LI > 0)$. More precisely, we ask whether for a given aperiodic system (X, T) we can find a positive monotone increasing sequence $\{a_n\}_{n=1}^\infty$ with $\lim_{n\to\infty} a_n = \infty$ such that (X, T) is of type $(LI > 0)$ for $(\{a_n\}_{n=1}^\infty, R(X))$. The following proposition shows that this is, in general, not the case. The proof of this proposition is simple, but rather laborious in its details; it can be found in [**1**].

PROPOSITION 3.12. *If $T : [0, 1) \to [0, 1)$ is the von Neumann–Kakutani adding machine and $\{a_n\}_{n=1}^\infty$ is a positive monotone increasing sequence with $\lim_{n\to\infty} a_n = \infty$, then there exists $\lambda \in [0, 1)$ such that for $I := [0, \lambda)$ we have*

$$\lim_{n\to\infty} \max\{m(A) \mid A \in \alpha(I)_0^{n-1}\} = 0 \quad \textit{and} \quad \liminf_{n\to\infty} \frac{H(\alpha(I)_0^{n-1})}{a_n} = 0.$$

We do not know whether there are completely ergodic systems for which results analogous to Proposition 3.12 are true.

Our second objective in §3 is to prove a pointwise analogue of Theorem 3.9. In fact, the statement of Theorem 3.13 is not exactly analogous to Theorem 3.9, because it says that every aperiodic system is pointwise of type $(LS \geq 1)$ for $(\{\log_2 n\}_{n=1}^\infty, R(X))$, and $\{\log_2 n\}_{n=1}^\infty$ is a stronger rate than $\{g(\log_2 n)\}_{n=1}^\infty$ in Theorem 3.9.

THEOREM 3.13. *If (X, T) is an aperiodic measure-preserving system and $E \in \mathcal{B}$ such that*

$$\lim_{n\to\infty} \max\{\mu(A) \mid A \in \alpha(E)_0^{n-1}\} = 0,$$

then

$$\limsup_{n\to\infty} \frac{I_{\alpha(E)_0^{n-1}}(x)}{\log_2 n} \geq 1 \quad \textit{a.e.}$$

In other words, (X, T) is pointwise of type $(LS \geq 1)$ for $(\{\log_2 n\}_{n=1}^\infty, R(X))$.

Proof. Let $M := \bigcap_{p=1}^\infty \bigcup_{n=1}^\infty X \setminus P_1^p(n,\ E)$. First we notice that $\mu(M) = 1$, because

$$\mu\left(\bigcup_{n=1}^\infty X \setminus P_1^p(n,\ E)\right) \geq \lim_{n\to\infty} \mu(X \setminus P_1^p(n,\ E)) = 1 - \lim_{n\to\infty} \mu(P_1^p(n,\ E)) = 1$$

for all $p \in \mathbb{N}$ (by Lemma 3.2). Now we wish to show that for every $N \in \mathbb{N}$ and $x \in M$, there exists an $n \geq N$ such that

$$p_n(s_n^E(x)) \geq \frac{n-1}{2}. \tag{1}$$

If $N \in \mathbb{N}$ and $x \in M$, then $x \in \bigcup_{n=1}^\infty P_1^N(n,\ E)$ and since $p_n(s_n^E(x)) \leq N$ for all $n \leq N$, we conclude that $X \setminus P_1^N(n,\ E) = \emptyset$ for all $n \leq N$. This shows that there must be an $n > N$ such that $x \in P_1^N(n-1,\ E)$ and $x \notin P_1^N(n,\ E)$ and, in particular, we have

$$p_n(s_n^E(x)) \neq p_{n-1}(s_{n-1}^E(x)). \tag{2}$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

Therefore, $p_n(s_n^E(x)) \geq (n-1)/2$, because if we had $p_n(s_n^E(x)) < (n-1)/2$, then Lemma 1.11 would imply that $s_n^E(x)$ had the same period as its subname $s_{n-1}^E(x)$, which would be a contradiction to (2). This completes the proof of (1). So for every $x \in M$, there exists a strictly increasing sequence $\{n_k\}_{k=1}^{\infty} \subset N$ such that

$$p_{n_k}(s_{n_k}^E(x)) \geq \frac{n_k - 1}{2} \quad \text{for all } k \in \mathbb{N}. \tag{3}$$

For every $k \in \mathbb{N}$ we choose a Rokhlin tower τ_k of length n_k, such that $\mu(\tau_k) > 1 - 1/n_k$. Then we obtain

$$\begin{aligned} \mu(A_{n_k}^E(s_{n_k}^E(x))) &\leq \mu(A_{n_k}^E(s_{n_k}^E(x)) \cap \tau_k) + \frac{1}{n_k} \\ &\leq \frac{2}{p_{n_k}(s_{n_k}^E(x))}\mu(\tau_k) + \frac{1}{n_k} \quad \text{(by 1.6)} \\ &\leq \frac{5}{n_k - 1} \quad \text{(by (3))}. \end{aligned} \tag{4}$$

Using the definition of the information function, we finally obtain

$$\liminf_{k\to\infty} \frac{I_{\alpha(E)_0^{n_k-1}}(x)}{\log_2 n_k} = \liminf_{k\to\infty} \frac{-\log_2 \mu(A_{n_k}^E(s_{n_k}^E(x)))}{\log_2 n_k} \geq 1 \quad \text{(by (4))}.$$

This shows that $\limsup_{n\to\infty} I_{\alpha(E)_0^{n-1}}(x)/\log_2 n \geq 1$ for all $x \in M$ and the proof is complete, because $\mu(M) = 1$. □

COROLLARY 3.14. *If (X, T) is ergodic and $E \in \mathcal{B}$ such that $\alpha(E)_0^{n-1} \neq \alpha(E)_0^n$ for all $n \in \mathbb{N}$, then*

$$\limsup_{n\to\infty} \frac{I_{\alpha(E)_0^{n-1}}(x)}{\log_2 n} \geq 1 \quad \textit{a.e.}$$

COROLLARY 3.15. *If (X, T) is completely ergodic and $E \in \mathcal{B}$ such that $0 < \mu(E) < 1$, then*

$$\limsup_{n\to\infty} \frac{I_{\alpha(E)_0^{n-1}}(x)}{\log_2 n} \geq 1 \quad \textit{a.e.}$$

In other words, every completely ergodic system is pointwise of type $(LS \geq 1)$ for $(\{\log_2 n\}_{n=1}^{\infty}, P(X))$.

The statement of Theorem 3.13 cannot, in general, be improved, as the following simple fact (proved in [**1**]) shows.

PROPOSITION 3.16. *Assume that $T : [0, 1) \to [0, 1)$ is the von Neumann–Kakutani adding machine and let $\alpha := \alpha([0, 1/3))$. Then*

$$\lim_{n\to\infty} \max\{\mu(A) \mid A \in \alpha_0^{n-1}\} = 0 \quad \textit{and} \quad \limsup_{n\to\infty} \frac{I_{\alpha_0^{n-1}}(x)}{\log_2 n} = 1 \quad \textit{for all } x \in [0, 1).$$

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

4. *Entropy convergence rates for rank-one transformations*

In this section we will mention two results for rank-one systems without going into any details of the proofs. Theorem 4.1 shows that the statement of Theorem 3.9 can be improved for rank-one mixing transformations, in the sense that any such transformation can be shown to be of type $(LS > 0)$ for $(\{\log_2 n\}_{n=1}^{\infty}, P(X))$. The idea for the proof of this theorem is as follows.

For a given $E \in \mathcal{B}$ with $0 < \mu(E) < 1$ and large $N \in \mathbb{N}$, we can choose a tower τ and a set F which is close to E in measure such that all the levels of τ are either contained in F or in $X \setminus F$. The goal is then to show that a sufficiently large subset of τ has 01-names with large periods, because then we can use Lemmas 1.6 and 1.8 as usual to find a lower bound for the entropy of $\alpha(E)_0^{n-1}$. So we make the counter assumption that a large subset of τ has 01-names with small periods and then construct a subtower σ of τ such that a large subset of the bottom level of σ has 01-names with small periods. The next step is to find a point in the bottom level of σ with a 01-name that has a small period p and differs from the 01-name s, generated by F, only in relatively few coordinates. In other words we show that s looks approximately like a 01-name with a small period. This implies that $\mu(E \cap T^{-p}E) \approx \mu(E) > \mu(E)^2$. But since (X, T) is assumed to be mixing, we have $\mu(E \cap T^{-p}E) \approx \mu(E)^2$, and this gives us a contradiction.

THEOREM 4.1. *If (X, T) is rank-one mixing and $E \in \mathcal{B}$ with $0 < \mu(E) < 1$, then*

$$\limsup_{n\to\infty} \frac{H(\alpha(E)_0^{n-1})}{\log_2 n} > 0.$$

Remark 4.2. It is an open question whether the result of Theorem 4.1 is still true if we weaken the assumptions on (X, T). We could, for example, replace 'mixing' by 'weakly mixing' or 'completely ergodic'. Furthermore, it is not clear whether the statement of Theorem 4.1 can be improved by showing that there is a $c > 0$ such that (X, T) is of type $(LS \geq c)$ for $(\{\log_2 n\}_{n=1}^{\infty}, P(X))$. We will see in [**2**] that a result like this can be shown for a certain class of weakly mixing rank-one systems.

The second result in this section shows that there are no rank-one systems with convergence rates of type $(LI \geq c)$ that are stronger than $\{\log_2 n\}_{n=1}^{\infty}$.

THEOREM 4.3. *If (X, T) is a rank-one system, then there is an $\alpha \in P(X)$ such that*

$$\liminf_{n\to\infty} \frac{H(\alpha_0^{n-1})}{\log_2 n} \leq 2.$$

Remark 4.4. The constructions in [**2**] show that the statement of Theorem 4.3 cannot be significantly improved except maybe for changes in the constant 2.

5. *Various other kinds of entropy convergence rates*

The preceding development could be varied in several ways.

Different entropy functions. The important properties of the entropy function $f(x) = -x \log_2 x$ which we used in §§2–4 are its concavity and, implicitly, the subadditivity

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

of the entropy of a finite partition, i.e. the inequality $H(\alpha \vee \beta) \leq H(\alpha) + H(\beta)$. The proof of this inequality uses the fact that $\log_2(ab) = \log_2 a + \log_2 b$. In other words, for $g(x) := f(x)/x$ we have $g(ab) = g(a) + g(b)$ for all $a, b \in (0, 1]$. It turns out that in order to prove subadditivity of H we only need the weaker assumption $g(ab) \leq g(a) + g(b)$ (see Lemma 5.2).

Definition 5.1. If $h : [0, 1] \to \mathbb{R}$ is a positive concave function with $h(0) = 0$ and α is a finite partition, then we define

$$H(\alpha)_h := \sum_{A \in \alpha} h(\mu(A)).$$

LEMMA 5.2. *Let $h : [0, 1] \to \mathbb{R}$ be a positive concave function with $h(0) = 0$ and assume that $g(x) := h(x)/x$ satisfies the inequality $g(ab) \leq g(a) + g(b)$ for all $a, b \in (0, 1]$. Then for all finite partitions α, β we have*

$$H(\alpha \vee \beta)_h \leq H(\alpha)_h + H(\beta)_h.$$

Proof.

$$\begin{aligned} H(\alpha \vee \beta)_h &\leq \sum_{\substack{A \in \alpha \\ B \in \beta}} \mu(A \cap B) \left(g\left(\frac{\mu(A \cap B)}{\mu(A)} \right) + g(\mu(A)) \right) \\ &= \sum_{\substack{A \in \alpha \\ B \in \beta}} \mu(A) h\left(\frac{\mu(A \cap B)}{\mu(A)} \right) + H(\alpha)_h \\ &\leq \sum_{B \in \beta} h\left(\sum_{A \in \alpha} \mu(A \cap B) \right) + H(\alpha)_h \quad \text{(because } h \text{ is concave)} \\ &= H(\beta)_h + H(\alpha)_h. \end{aligned}$$

□

It is easy to check that, for example, the function $h(x) := x \log_2(1 - \log_2 x)$ satisfies the assumptions of Lemma 5.2. For any such h we can define generalized entropy convergence rates in the obvious way.

We say that a measure-preserving system (X, T) is of type $(LS \geq c)_h$ for $(\{a_n\}_{n=1}^{\infty}, P)$ if

$$\limsup_{n \to \infty} \frac{H(\alpha_0^{n-1})_h}{a_n} \geq c \quad \text{for all } \alpha \in P.$$

Similarly, we can generalize the definitions of all the other convergence types. It should be possible to restate most of our results in terms of these generalized concepts by making some minor adjustments in the arguments. For example, for $h(x) = x \log_2(1 - \log_2 x)$ we should be able to show that every aperiodic system is of type $(LS = \infty)_h$ for $(\{g(\log_2(\log_2 x))\}_{n=1}^{\infty}, R(X))$ if $g : [0, \infty) \to \mathbb{R}$ is a positive monotone increasing function with $\int_1^{\infty} g(x)/x^2 \, dx < \infty$. This would be a generalization of Theorem 3.9. For the proof of this statement we could still use Lemma 3.4, with the only difference being that the x_k's would not be defined to be basically the measure of the set of all points in X which have 01-names with periods ranging between 2^{k-1} and $2^k - 1$ (see the

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

proof of Proposition 3.8), but instead we would define x_k to be the measure of the set of all points in X which have 01-names with periods ranging between $2^{2^{k-1}}$ and $2^{2^k} - 1$.

The most interesting question is of course whether these generalized convergence rates are different from the rates defined via the ordinary entropy function $f(x) = -x \log_2 x$. We believe that this is indeed the case. More precisely, by using the techniques developed in [**2**], it should be possible to construct systems which have the same convergence rates if we use $f(x) = -x \log_2 x$ but have different rates for $h(x) = x \log_2(1 - \log_2 x)$, and vice versa. So we would obtain a large set of essentially different isomorphism invariants corresponding to the class of functions h satisfying the assumptions of Lemma 5.2.

Even further generalizations seem possible, because in many (or maybe all) arguments the subadditivity of H is apparently not really necessary but merely convenient. So it may be possible to obtain further invariants if we allow entropy functions like $h(x) := \sqrt{x}$ for which $g(x) := h(x)/x$ does not satisfy the inequality $g(ab) \leq g(a) + g(b)$ for all $a, b \in (0, 1]$.

Counting-type convergence rates. Another variation of entropy convergence rates is related to an existing concept in the theory of entropy. Let $n \in \mathbb{N}$ and $\varepsilon > 0$. If (X, T) is a measure-preserving system and α is a finite partition of X, then we define $N(T, \alpha, n, \varepsilon)$ to be the minimal number of atoms in α_0^{n-1} which cover a subset of X with measure greater than $1 - \varepsilon$. It can be shown (see [**8**]) that if T is ergodic, then

$$h(T, \alpha) = \lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{\log_2(N(T, \alpha, n, \varepsilon))}{n},$$

where $h(T, \alpha)$ denotes the entropy of T with respect to α. The existence of the limit as $\varepsilon \to 0$ is obvious, because $N(T, \alpha, n, \varepsilon)$ is decreasing in ε. This means that the definition of entropy can be based on the counting of names, and we are thus led to the following counting-type convergence rates. We say that (X, T) is of type C–$(LS \geq c)$ for $(\{a_n\}_{n=1}^{\infty}, P)$ if

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log_2(N(T, \alpha, n, \varepsilon))}{a_n} \geq c \quad \text{for all } \alpha \in P,$$

and (X, T) is of type C–$(LI \geq c)$ for $(\{a_n\}_{n=1}^{\infty}, P)$ if

$$\lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{\log_2(N(T, \alpha, n, \varepsilon))}{a_n} \geq c \quad \text{for all } \alpha \in P.$$

Similarly, all the other convergence types ($(LS \leq c)$, $(LI \leq c)$, $(LS = \infty)$ etc) can be defined as counting types. It is clear that all these types are isomorphism invariants. Furthermore, it is conceivable that, in contrast to the ordinary types $(LS \leq c)$ and $(LI \leq c)$, the counting types C–$(LS \leq c)$ and C–$(LI \leq c)$ can be used to distinguish systems with entropy zero. The reason for this conjecture is that the proof of Theorem 2.8 was based on the construction of independent sets in a sequence of Rokhlin towers τ_n, where the measure of these towers was decreasing to zero at an exponential rate. An argument like this cannot be successful for counting types, because if the measure of τ_n is eventually smaller than ε, then the atoms in τ_n are not necessarily making a contribution to the value of $N(T, \alpha, n, \varepsilon)$. Although this seems to be an advantage of the counting

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

types, it may be more difficult to work with them, because $N(T, \alpha, n, \varepsilon)$ might be harder to calculate than $H(\alpha_0^{n-1})$. As in the case of ordinary convergence types, a change in the entropy function might lead to a whole variety of different counting types. We could, for example, replace the function $\log_2 x$ in the definition of the counting types by $\log_2(1 - \log_2 x)$.

Equivalence-type convergence rates. The final variation opens a connection to the theory of Kakutani equivalence (see [**3**], [**5**] or [**6**]). The definitions are based on the $\bar{f}$-metric, which is defined as follows.

Definition 5.3. For $n, N \in \mathbb{N}$ and $s, r \in \{1, \ldots, N\}^n$ let

$$\bar{f}_n(s, r) := 1 - \frac{k}{n},$$

where k is the largest integer for which we can find subsequences $i_1 < \cdots < i_k$, $j_1 < \cdots < j_k$ with

$$s(i_m) = r(j_m) \quad \text{for all } m \in \{1, \ldots, k\}.$$

Considering again a measure-preserving system (X, T), we denote by $PF(X)$ the set of all finite partitions of X. If $\alpha = \{A_1, \ldots, A_N\} \in PF(X)$ and $n \in \mathbb{N}$, then we define, in analogy to Definition 1.3,

$$p_\alpha(x) := i \quad \text{for all } x \in A_i \quad \text{and} \quad s_n^\alpha(x) := (p_\alpha(x), p_\alpha(Tx), \ldots, p_\alpha(T^{n-1}x)).$$

Furthermore, for $\varepsilon > 0$ and $x \in X$ we define the ε-ball around x with respect to $\bar{f}_n$ and α as

$$B_n^\alpha(x, \varepsilon) := \{y \in X \mid \bar{f}_n(s_n^\alpha(x), s_n^\alpha(y)) < \varepsilon\}$$

and by $M(T, \alpha, n, \varepsilon)$ we denote the minimal number of balls $B_n^\alpha(x, \varepsilon)$ which cover a subset of X with measure greater than $1 - \varepsilon$. Using this notation, we define the equivalence-type convergence rates as follows. We say that (X, T) is of type $E\text{–}(LS \geq c)$ for $\{a_n\}_{n=1}^\infty$ if

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log_2(M(T, \alpha, n, \varepsilon))}{a_n} \geq c \quad \text{for all } \alpha \in PF(X)$$

and (X, T) is of type $E\text{–}(LI \geq c)$ for $\{a_n\}_{n=1}^\infty$ if

$$\lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{\log_2(M(T, \alpha, n, \varepsilon))}{a_n} \geq c \quad \text{for all } \alpha \in PF(X).$$

Similarly, we can define all the other convergence types ($(LS \leq c)$, $(LI \leq c)$, $(LS = \infty)$ etc) as equivalence types. The existence of the limit as $\varepsilon \to 0$ is again obvious. It is also clear that these convergence types are isomorphism invariants, and it is not too difficult to show that the types with an upper estimate (i.e. $E\text{–}(LS \leq c)$, $E\text{–}(LI \leq c)$, $E\text{–}(LS < \infty)$ and $E\text{–}(LI < \infty)$) are invariants of Kakutani equivalence. Surprisingly, the equivalence types with a lower estimate are apparently not invariants of Kakutani equivalence, or at least it is more difficult to show the invariance in this case. The usefulness of equivalence type convergence rates has already been demonstrated by Marina Ratner, who in [**7**]

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press

used a generalization of these types to the case of measure-preserving flows in order to prove that different cartesian powers of the horocycle flow are inequivalent. Interesting generalizations may be obtained by changing the entropy function or by replacing $\bar{f}_n$ by other metrics like, for example, the Hamming metric (see [**5**]). In the latter case, though, the convergence rates would no longer be invariants of Kakutani equivalence. Interesting connections might also exist between equivalence-type convergence rates and the theory of restricted orbit equivalence as presented in [**9**].

Acknowledgements. I am very thankful to Dr Karl E. Petersen for many helpful discussions and his constant encouragements. I am also grateful to Dr Anatole Katok and Dr J.-P. Thouvenot for pointing out a connection of my research to the work of Dr Marina Ratner.

This research was supported in part by NSF Grant DMS-9203489.

REFERENCES

[1] F. Blume. The rate of entropy convergence. *Doctoral Dissertation*, University of North Carolina at Chapel Hill, 1995.
[2] F. Blume. The rate of entropy convergence as an isomorphism invariant. To appear.
[3] J. Feldman. New K-automorphisms and a problem of Kakutani. *Israel J. Math.* **38** (1976), 16–38.
[4] S. Ferenczi. Systèmes de rang fini. *Thèse de Doctorat*, Universite d' Aix-Marseille 2, 1990.
[5] A. B. Katok. Monotone equivalence in ergodic theory. *Math. USSR Izvestija* **11** (1977), 99–146.
[6] D. S. Ornstein, D. J. Rudolph and B. Weiss. Equivalence of measure preserving transformations. *Mem. Amer. Math. Soc.* **37** (1982).
[7] M. Ratner. Some invariants of Kakutani equivalence. *Israel J. Math.* **38** (1981), 231–240.
[8] D. J. Rudolph. *Fundamentals of Measurable Dynamics*. Oxford University Press, New York, 1990.
[9] D. J. Rudolph. Restricted orbit equivalence. *Mem. Amer. Math. Soc.* **54** (1985).
[10] P. C. Shields. The ergodic theory of discrete sample paths. Preliminary version, 1994.

https://doi.org/10.1017/S0143385797069733 Published online by Cambridge University Press