Can. J. Math.Vol. 44 (1), 1992 pp. 215–224

SEQUENCE ENTROPY AND MILD MIXING

QING ZHANG

0. **Introduction.** Entropy characterizations of different spectral and mixing properties of dynamical systems were dealt with by a number of authors (see [5], [6] and [8]).

Given an infinite subset $\Gamma = \{t_n\}$ of $\mathbf{N}$ and a dynamical system $(\mathbf{X}, \mathcal{B}, \mu, T)$ one can define sequence entropy along Γ: $h_\Gamma(T, \xi) = \overline{\lim}_n \frac{1}{n} H\big(\bigvee_{i=1}^n T^{t_i}\xi\big)$ for any finite partition ξ, and $h_\Gamma(T) = \sup_\xi h_\Gamma(T, \xi)$. In [6] Kushnirenko used sequence entropy to give a characterization of systems with discrete spectrum.

THEOREM 0.1 (A. G. KUSHNIRENKO [6]). *An invertible measure preserving transformation T has discrete spectrum if and only if $h_\Gamma(T) = 0$ for all $\Gamma \subset \mathbf{N}$.*

In [8] A. Saleski gave a characterization of weak mixing and later Hulse [4] improved Saleski's result. The following theorem is slightly stronger than theirs. For its proof, see Appendix.

THEOREM 0.2. *An invertible measure preserving transformation T is weakly mixing if and only if for any set $\Gamma \subset \mathbf{N}$ with positive density, there is a subset Γ_1 of Γ such that for any finite partition ξ, $h_{\Gamma_1}(T, \xi) = H(\xi)$.*

In [3] H. Furstenberg and B. Weiss introduced a new kind of mixing property of dynamical system which they called mild mixing. A function $f \in L^2(\mathbf{X}, \mathcal{B}, \mu)$ is rigid if there exists $\{t_n\}$ such that $T^{t_n}f \longrightarrow f$ in L^2-topology. A transformation T is rigid if there is $\{t_n\}$ such that for any $f \in L^2(\mathbf{X}, \mathcal{B}, \mu)$, $T^{t_n}f \longrightarrow f$ in L^2-topology. A transformation T is mild mixing if there is no nonconstant rigid functions in $L^2(\mathbf{X}, \mathcal{B}, \mu)$. Mild mixing is not "weaker" than weak mixing since any eigenfunction is also a rigid function. It follows from the definition that mild mixing is not " stronger" than strong mixing. In fact weak mixing is really "weaker" and strong mixing is really "stronger" than mild mxing. The reader can find details in [3]. The purpose of this note is to give the following sequence entropy characterizations of rigidity and mild mixing.

THEOREM 0.3. *Let T be an invertible measure preserving transformation. Then T is rigid if and only if there exists a subset Γ of $\mathbf{N}$ such that if $\{F_i\}$ is any sequence of pairwise disjoint finite subsets of Γ and $s_i = \sum_{a \in F_i} a$, then $h_{\{s_i\}}(T) = 0$.*

This paper forms a part of the author's Ph.D. thesis under the direction of Professor V. Bergelson at The Ohio State University. The author wishes to thank him for his continued help and guidance.

Received by the editors May 14, 1990.

AMS subject classification: 28D20.

THEOREM 0.4. *T is mild mixing if and only if for any subset Γ of $\mathbf{N}$, there is a sequence $\{F_n\}$ of pairwise disjoint finite subsets of Γ such that for any finite partition ξ of $\mathbf{X}$, $h_{\{s_i\}}(T,\xi) = H(\xi)$, where $s_i = \sum_{a\in F_i} a$.*

1. **Preliminaries.** In this paper, we will be dealing with a probability measure space $(\mathbf{X}, \mathcal{B}, \mu)$ together with a one-to-one transformation $T: \mathbf{X} \longrightarrow \mathbf{X}$ which is measure preserving, i.e. $\mu(B) = \mu(T^{-1}B)$ for $B \in \mathcal{B}$. A partition ξ of $\mathbf{X}$ is a disjoint collection of elements of $\mathcal{B}$ whose union is $\mathbf{X}$. $\#(\xi)$ will be used to denote the number of elements in ξ. If $\#(\xi) = k$, we will call ξ a k-cell partition. Suppose ξ and η are two finite partitions of $\mathbf{X}$. We write $\xi \leq \eta$ to mean that each element of ξ is a union of elements of η.

DEFINITION 1.1. Suppose ξ and η are two finite partitions of $\mathbf{X}$. The entropy of ξ, written $H(\xi)$, is defined by the formula

$$H(\xi) = -\sum_{A\in\xi} \mu(A)\ln\mu(A).$$

The entropy of ξ given η, written $H(\xi \mid \eta)$, is defined by the formula:

$$H(\xi \mid \eta) = -\sum_{B\in\eta}\sum_{A\in\xi} \mu(A)\ln\frac{\mu(A\cap B)}{\mu(B)}$$

DEFINITION 1.2 (KUSHNIRENKO [6]). Suppose ξ is a finite partition of $\mathbf{X}$ and $\Gamma = \{t_n\}$ is an infinite sequence of positive integers. Define:

$$h_\Gamma(T,\xi) = \overline{\lim_n}\frac{1}{n}H\Big(\bigvee_{i=1}^{n} T^{t_i}\xi\Big) \text{ and } h_\Gamma(T) = \sup_\xi h_\Gamma(T,\xi).$$

$h_\Gamma(T)$ is called *sequence entropy* along Γ (or Γ-entropy). In particular, if $\Gamma = \mathbf{N}$, we will call it *entropy of* T and denote it by $h(T)$.

We will say that two ordered k-cell partitions $\xi = \{A_1,\ldots,A_k\}$ and $\eta = \{B_1,\ldots,B_k\}$ of $\mathbf{X}$ are equivalent if $\mu(A_i\Delta B_i) = 0$ for $i = 1,2,\ldots,k$. Let Z_k be the set of equivalence classes of ordered k-cell partitions. Define a complete metric on $\mathbf{Z}_k$ by:

$$|\xi - \eta| = \sum_{i=1}^{k}\mu(A_i\Delta B_i).$$

A subset $K \subset \mathbf{Z}_k$ is totally bounded if it is totally bounded in the metric defined above. It is easy to see that K is totally bounded if and only if for any $\varepsilon > 0$ there is a finite subset F of K such that $\inf_{\eta\in F}|\xi - \eta| < \varepsilon$ for all $\xi \in K$.

We use $\mathbf{N}$,$\mathbf{Z}$ and $\mathcal{F}$ to denote respectively the set of all positive integers, the set of all integers and the set of all finite nonempty subsets of $\mathbf{N}$.

The following definitions mainly come from [1], [2] and [3].

DEFINITION 1.3 (CF. [1], [2]). A set of positive integers is called an IP-set if there exists a sequence $p_1, p_2, \ldots$ such that the set in question consists of the numbers p_i together with all finite sums $p_{i_1} + p_{i_2} + \cdots + p_{i_k}$ with $i_1 < i_2 < \cdots < i_k$.

DEFINITION 1.4 (CF. [1]). A homomorphism $\psi\colon \mathcal{F} \longrightarrow \mathcal{F}$ is a map such that $\alpha \cap \beta = \emptyset$ implies $\psi(\alpha) \cap \psi(\beta) = \emptyset$ and $\psi(\alpha \cup \beta) = \psi(\alpha) \cup \psi(\beta)$.

An $\mathcal{F}$-sequence is a sequence $\{x_\alpha; \alpha \in \mathcal{F}\}$ indexed by elements $\alpha \in \mathcal{F}$. Given a semigroup $\mathbf{M}$ and a sequence $\{x_i\}$ of elements of $\mathbf{M}$, one can define an $\mathcal{F}$-sequence by

$$x_{\{i_1,i_2,\ldots,i_k\}} = x_{i_1}x_{i_2}\ldots x_{i_k}$$

where $i_1 < i_2 < \cdots < i_k$. Such an $\mathcal{F}$-sequence will be called an IP-system. Given an $\mathcal{F}$-sequence $\{x_\alpha\}$ and a homomorphism $\psi\colon \mathcal{F} \longrightarrow \mathcal{F}$, one can define $\mathcal{F}$-subsequence $\{y_\alpha = x_{\psi(\alpha)}\}$. In particular, if $\{x_\alpha\}$ is an IP-system, we will call $\{y_\alpha\}$ a sub-IP-system. If $(\mathbf{X}, \mathcal{B}, \mu, T)$ is a dynamical system and $\Gamma = \{t_n\} \subset \mathbf{N}$, then the IP-system of measure preserving transformations generated by Γ is defined by $\Sigma_\Gamma = \{T_\alpha = \prod_{i\in\alpha} T^{t_i}; \alpha \in \mathcal{F}\}$.

DEFINITION 1.5 (CF. [1],[2] AND [3]). Let $\{x_\alpha\}$ be an $\mathcal{F}$-sequence in a topological space $\mathbf{M}$ and $x \in \mathbf{M}$. We say that x is IP-limit of $\{x_\alpha\}$ if for every neighborhood V of x there exists an index β such that $\alpha \cap \beta = \emptyset$ implies $x_\alpha \in V$.

REMARK. If $\mathbf{M}$ is a Hausdorff topological space, the IP-limit is unique. All the topological spaces we deal with in this paper are Haudorff.

DEFINITION 1.6 (CF. [2] AND [3]). A function $f \in L^2(\mathbf{X}, \mathcal{B}, \mu)$ is called *rigid* if there is an infinite subset $\{t_n\}$ of $\mathbf{N}$ such that $\lim_n T^{t_n}f = f$ in L_2-norm. A measure preserving transformation T is called *rigid* if there is an infinite subset $\{t_n\}$ of $\mathbf{N}$ such that $\lim_n T^{t_n}f = f$ in L_2-norm for all $f \in L^2(\mathbf{X}, \mathcal{B}, \mu)$.

REMARK. An equivalent definition for rigid transformation is: there is an infinite subset $\Gamma \subset \mathbf{N}$ such that $\mathrm{IP}-\lim_{T_\alpha\in\Sigma_\Gamma} T_\alpha f = f$ in L_2-norm for all $f \in L^2(\mathbf{X}, \mathcal{B}, \mu)$. For details, see [1, p.141].

DEFINITION 1.7. A dynamical system $(\mathbf{X}, \mathcal{B}, \mu, T)$ is called *mildly mixing* if it has no nonconstant rigid functions.

For the proofs of the following two theorems see [1, p. 140,145].

THEOREM 1.1. *Suppose $(\mathbf{X}, \mathcal{B}, \mu, T)$ is a dynamical system. Then for any subset Γ of $\mathbf{N}$ there is an IP-subsystem Σ' of Σ_Γ and an orthogonal projection $\mathbf{P}$ such that:*

$$\mathrm{IP}-\lim_{T_\alpha\in\Sigma'}\langle T_\alpha f, g\rangle = \langle \mathbf{P}f, g\rangle \quad f, g \in L^2(\mathbf{X}, \mathcal{B}, \mu).$$

In particular, if $\mathbf{P}f = f$, then:

$$\mathrm{IP}-\lim_{T_\alpha\in\Sigma'} \|T_\alpha f - f\|_2 = 0.$$

THEOREM 1.2. *T is mildly mixing if and only if for any infinite subset Γ of $\mathbf{N}$, there exists an IP-subsystem Σ' of Σ_Γ such that:*

$$\mathrm{IP}-\lim_{T_\alpha\in\Sigma'}\langle T_\alpha f, g\rangle = \langle f, 1\rangle\langle 1, g\rangle \quad f, g \in L^2(\mathbf{X}, \mathcal{B}, \mu).$$

PROPOSITION 1.3. *Let* $\mathbf{P}$ *be the projection defined as above. There is a sub-σ-algebra* $\mathcal{B}_0$ *such that*

$$\mathbf{P}\big(L^2(\mathbf{X}, \mathcal{B}, \mu)\big) = L^2(\mathbf{X}, \mathcal{B}_0, \mu).$$

PROOF. Let $H_0 = \mathbf{P}\big(L^2(\mathbf{X}, \mathcal{B}, \mu)\big)$ and let $\mathcal{B}_0$ be the smallest sub-σ-algebra with respect to which the functions in H_0 are measurable. Then $L^2(\mathbf{X}, \mathcal{B}_0, \mu) \supset H_0$. It is obvious that all of constants are in H_0. Since $\big\| T_\alpha |f| - |f| \big\|_2 \leq \| T_\alpha f - f \|_2$, we have that $f \in H_0$ implies $|f| \in H_0$ (notice that $f \in H_0$ if and only if $\mathrm{IP} - \lim_{T_\alpha \in \Sigma'} T_\alpha f = f$). So if $f, g \in H_0$,

$$\max\{f, g\} = \frac{f + g + |f - g|}{2} \in H_0; \quad \min\{f, g\} = \frac{f + g - |f - g|}{2} \in H_0.$$

If $f \in H_0$ and L is a real number, then:

$$\mathbf{1}_{\{x; f(x) > L\}} = \lim_n \min\big\{ n(\max\{f, L\} - L), 1 \big\} \in H_0.$$

Suppose $\mathbf{1}_A, \mathbf{1}_B \in H_0$. Since:

$$\| T_\alpha \mathbf{1}_{A \cap B} - \mathbf{1}_{A \cap B} \|_2 \leq \| T_\alpha \mathbf{1}_A - \mathbf{1}_A \|_2 + \| T_\alpha \mathbf{1}_B - \mathbf{1}_B \|_2$$

we have $\mathbf{1}_{A \cap B} \in H_0$. So $H_0 \supset L^2(\mathbf{X}, \mathcal{B}_0, \mu)$.

2. **Entropy characterization of mild mixing and rigidity.** The following two lemmas are generalizations of Proposition and Lemma 3 in [6].

LEMMA 2.1. *For every* $\varepsilon > 0$ *and positive integer* p *there is* $\delta = \delta(\varepsilon, p)$ *such that for any ordered* p*-cell partition* η *and* q*-cell partition* ξ *with* $p \leq q$ *the inequality* $H(\eta \mid \xi) < \delta$ *implies the existence of an ordered* p*-cell partition* $\xi' \leq \xi$ *such that* $|\eta - \xi'| < \varepsilon$.

PROOF. For $\varepsilon > 0$ we choose $L > 1$, $0 < \delta' < 1/2$ and $0 < \delta < 1/2$ such that:

(1) $(p + 1)(\delta' + 1/L) < \varepsilon$ and $L\delta < 1/2$.

(2) for $0 \leq y \leq 1$, $-y \ln y < L\delta$ implies $y < \delta'$ or $1 - y < \delta'$.

Now we assume that $\eta = \{A_1, \ldots, A_p\}$, $\xi = \{B_1, \ldots, B_q\}$ are two finite partitions and $H(\eta \mid \xi) < \delta$. Define:

$$\mu_j(A) = \frac{\mu(A \cap B_j)}{\mu(B_j)} \text{ for } j = 1, 2, \ldots q$$

and

$$H_j(\eta) = -\sum_{i=1}^{p} \mu_j(A_i) \ln \mu_j(A_i) \text{ for } j = 1, 2, \ldots q.$$

Suppose $H_j(\eta) < L\delta$ for $1 \leq j \leq q_0$ and $H_j(\eta) \geq L\delta$ for $q_0 + 1 \leq j \leq q$. Since $H(\eta \mid \xi) = \sum_{j=1}^{q} H_j(\eta) \mu(B_j) < \delta$, we have $\sum_{j=q_0+1}^{q} \mu(B_j) < \frac{1}{L}$. Let $C_0 = \bigcup_{j=q_0+1}^{q} B_j$. It follows from the definition that $H_j(\eta) < L\delta$ implies $-\mu_j(A_i) \ln \mu_j(A_i) < L\delta$. Then by (2)

$\mu_j(A_i) < \delta'$ or $1 - \mu_j(A_i) < \delta'$ for $1 \leq j \leq q_0$ and $1 \leq i \leq p$. Since μ_j is a probability measure, there is a unique A_i such that $1 - \mu_j(A_i) < \delta'$. Hence we can assign A_i to B_j. Now we have a map τ defined by above assignment from $B_1, \ldots, B_{q_0}$ to $A_1, \ldots, A_p$. Let $C_i = \bigcup_{\tau(B_j)=A_i} B_j$ (if there is not B_j satisfying $\tau(B_j) = A_i$, take C_i to be the empty set). Then:

$$\mu(C_i \cap A_i) = \sum_{\tau(B_j)=A_i} \mu(B_j \cap A_i) > (1-\delta') \sum_{\tau(B_j)=A_i} \mu(B_j) = (1-\delta')\mu(C_i).$$

This means $\mu(C_i \cap A_i^c) < \delta'\mu(C_i)$. On the other hand:

$$\mu(A_i \cap C_i^c) \leq \sum_{\tau(B_j)\neq A_i} \mu(A_i \cap B_j) + \mu(C_0) < \sum_{\tau(B_j)\neq A_i} \delta'\mu(B_j) + \frac{1}{L} \leq \delta' + \frac{1}{L}.$$

Now taking the partition $\xi' = \{C_0 \cup C_1, C_2, \ldots, C_p\}$, we get:

$$|\eta - \xi'| \leq \delta' + p\Big(\delta' + \frac{1}{L}\Big) + \frac{1}{L} < \varepsilon.$$

LEMMA 2.2. *Suppose $\xi = \{A_1, \ldots A_k\}$ is a partition of $\mathbf{X}$ and $\Gamma = \{s_n\} \subset \mathbf{N}$. If for any $\Gamma_1 \subset \Gamma$, $h_{\Gamma_1}(T,\xi) = 0$ then $\{T^{s_n}\xi\}$ is a totally bounded set.*

PROOF. If $\{T^{s_n}\xi\}$ is not a totally bounded set, without loss of generality we suppose that there exists an $\varepsilon_0 > 0$ such that $|T^{s_m}\xi - T^{s_n}\xi| \geq \varepsilon_0$. By Lemma 2.1 we can find a $\delta = \delta\big(\frac{\varepsilon_0}{2}, k\big)$. Now we inductively construct $\Gamma_1 = \{t_n\} \subset \Gamma$ such that:

$$H\Big(T^{t_n}\xi \mid \bigvee_{j=1}^{n-1} T^{t_j}\xi\Big) \geq \delta.$$

Suppose we already have $t_1 < \cdots < t_{n-1}$. If for any $s_m > t_{n-1}$

$$H\Big(T^{s_m}\xi \mid \bigvee_{j=1}^{n-1} T^{t_j}\xi\Big) < \delta$$

then by Lemma 2.1 there is a k-cells partition $\eta_m < \bigvee_{j=1}^{n-1} T^{t_j}\xi$ such that $|T^{s_m}\xi - \eta_m| < \frac{\varepsilon_0}{2}$. Since $\bigvee_{j=1}^{n-1} T^{t_j}\xi$ is a finite partition, there must be $\eta_{m_1} = \eta_{m_2}$ with $m_1 \neq m_2$. This implies $|T^{s_{m_1}}\xi - T^{s_{m_2}}\xi| < \varepsilon_0$. But $|T^{s_m}\xi - T^{s_n}\xi| \geq \varepsilon_0$. So there exists s_m such that:

$$H\Big(T^{s_m}\xi \mid \bigvee_{j=1}^{n-1} T^{t_j}\xi\Big) \geq \delta.$$

Put $t_n = s_m$. Continuing in this fashion, we get a sequence $\Gamma_1 = \{t_n\}$ such that:

$$h_{\Gamma_1}(T,\xi) = \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{j=1}^{n} H\Big(T^{t_j}\xi \mid \bigvee_{i=1}^{j-1} T^{t_i}\xi\Big) \geq \delta.$$

This contradicts our assumption that for any Γ_1 contained in Γ one has $h_{\Gamma_1}(T,\xi) = 0$.

LEMMA 2.3. *For any $\varepsilon > 0$ and integer k, there is a $\delta = \delta(\varepsilon, k)$ (which depends only on ε and k) such that if k-cell partitions ξ and η satisfy $|\xi - \eta| < \delta$, then $\big|H(\xi \mid \eta) + H(\eta \mid \xi)\big| < \varepsilon$.*

For a proof, see Lemma 4.15 in [9].

THEOREM 0.3. *T is rigid if and only if there exists a subset Γ of $\mathbf{N}$ such that if $\{F_i\}$ is any sequence of pairwise disjoint finite subsets of Γ and $s_i = \sum_{a\in F_i} a$, then $h_{\{s_i\}}(T) = 0$.*

PROOF. Suppose T is not rigid. Then for any $\Gamma = \{a_n\}$, there is an IP-subsystem Σ' of Σ_Γ such that the range of the projection $\mathbf{P}$, defined by $\langle \mathbf{P}f, g\rangle = \mathrm{IP}-\lim_{T_\alpha\in\Sigma'}\langle T_\alpha f, g\rangle$, is not the whole space $L^2(\mathbf{X}, \mathcal{B}, \mu)$. By Proposition 1.3, we have a sub-σ-algebra $\mathcal{B}_0$ such that

$$\mathbf{P}\big(L^2(\mathbf{X}, \mathcal{B}, \mu)\big) = L^2(\mathbf{X}, \mathcal{B}_0, \mu).$$

Then there is a $B' \in \mathcal{B}$ such that $B' \notin \mathcal{B}_0$.

Take $D = \{x; \varepsilon < \mathbb{E}(\mathbf{1}_{B'} \mid \mathcal{B}_0) < 1 - \varepsilon\} \in \mathcal{B}_0$. Then $\mu(D) > 0$ for some $\varepsilon > 0$. (Otherwise $\mathbb{E}(\mathbf{1}_{B'} \mid \mathcal{B}_0) = 1$ or 0 which means $B' \in \mathcal{B}_0$). Let $B = B' \cap D$. Then $\varepsilon < \mathbb{E}(\mathbf{1}_B \mid \mathcal{B}_0) < 1 - \varepsilon$ on D and $\mathbb{E}(\mathbf{1}_B \mid \mathcal{B}_0) = 0$ on D^c. Take $\xi = \{B, B^c\}$. Our next step is to find a sequence $\{\alpha_n\}$ of pairwise disjoint finite subsets of $\mathbf{N}$ such that for $t_n = \sum_{i\in\alpha_n} a_i\{T^{t_n}\xi\}$ is not totally bounded.

Suppose $\alpha_1, \ldots, \alpha_{n-1}$ have been chosen and:

$$\mu(T_{\alpha_i}B \cap T_{\alpha_j}B) < \Big(1 - \frac{1}{2}\varepsilon\Big)\mu(B)$$

for all $0 \le i < j \le n-1$.

Let $f_B = \mathbf{1}_B - \mathbb{E}(\mathbf{1}_B \mid \mathcal{B}_0)$. By Theorem 1.1 we have:

$$\mathrm{IP}-\lim_{T_\alpha\in\Sigma'} \int T_\alpha f_B \mathbf{1}_C \, d\mu = 0$$

for all $C \in \mathcal{B}_0$. Now we choose α_n such that:

(1) $\alpha_n \cap \alpha_i = \emptyset \quad i = 1, 2, \ldots$.

(2) $\int T_{\alpha_n} f T_{\alpha_j}\mathbf{1}_B \, d\mu \le \frac{1}{2}\varepsilon\mu(B)$ for all $j \le n-1$.

Then:

$$\begin{aligned}\mu(T_{\alpha_n}B \cap T_{\alpha_j}B) &= \int T_{\alpha_n}\mathbf{1}_B T_{\alpha_j}\mathbf{1}_B \, d\mu\\ &= \int T_{\alpha_n}\mathbb{E}(\mathbf{1}_B \mid \mathcal{B}_0)T_{\alpha_j}\mathbf{1}_B \, d\mu + \int T_{\alpha_n} f_B T_{\alpha_j}\mathbf{1}_B \, d\mu\\ &\le (1-\varepsilon)\mu(B) + (1/2)\varepsilon\mu(B) = \{1 - (1/2)\varepsilon\}\mu(B)\end{aligned}$$

for all $0 \le j \le n-1$. Now we already have $\{\alpha_n\}$. Hence

$$\begin{aligned}\mu(T_{\alpha_n}B \Delta T_{\alpha_j}B) &= \int (T_{\alpha_n}\mathbf{1}_B - T_{\alpha_j}\mathbf{1}_B)^2 \, d\mu\\ &= 2\mu(B) - 2\int T_{\alpha_n}\mathbf{1}_B T_{\alpha_j}\mathbf{1}_B \, d\mu \ge \varepsilon\mu(B) > 0.\end{aligned}$$

This implies $\{T_{\alpha_n}\xi\}$ is not totally bounded. By Lemma 2.2 there exists $\Gamma_1 \subset \{t_n = \sum_{i\in\alpha_n} a_i\}$ such that $h_{\Gamma_1}(T,\xi) > 0$.

Conversely, if T is rigid, there is a subset $\Gamma = \{s_n\}$ of $\mathbf{N}$ such that $\mathrm{IP}-\lim_{T_\alpha\in\Sigma_\Gamma} T_\alpha \mathbf{1}_A = \mathbf{1}_A$ for all $A \in \mathcal{B}$. Then for a sequence $\{\alpha_n\}$ of pairwise disjoint finite subsets of $\mathbf{N}$ $\lim_{n\to\infty} T_{\alpha_n}\mathbf{1}_A = \mathbf{1}_A$. This implies $\lim_{n\to\infty} H(T_{\alpha_n}\xi \mid \xi) = 0$ for any finite partition ξ. Taking $t_n = \sum_{i\in\alpha_n} s_i$, we have:

$$h_{\{t_n\}}(T,\xi) = \lim_{n\to\infty}\frac{1}{n}H\Big(\bigvee_{j=0}^{n-1}T^{t_j}\xi\Big) = \lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}H\Big(T^{t_j}\xi \mid \bigvee_{i=0}^{j-1}T^{t_i}\xi\Big)$$
$$\leq \lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}H(T^{t_j}\xi \mid \xi) = 0.$$

This proves our theorem.

LEMMA 2.4. *For any Lebesgue space $(\mathbf{X},\mathcal{B},\mu)$, there is a countable set $\{\xi_k\}$ of finite partitions such that for any finite partition ξ,* $\inf_k\{H(\xi \mid \xi_k) + H(\xi_k \mid \xi)\} = 0$.

This lemma is an immediate corollary of 6.3, [7].

THEOREM 0.4. *T is mild mixing if and only if for any subset Γ of positive integers, there is a sequence $\{F_i\}$ of pairwise disjoint finite subsets of Γ such that for any finite partition ξ of $\mathbf{X}$, $h_{\{s_i\}}(T,\xi) = H(\xi)$, where $s_i = \sum_{a\in F_i} a$.*

PROOF. Suppose T is mild mixing. For $\Gamma = \{a_i\}$, we define an IP-system $\Sigma_\Gamma = \{T_\alpha = \prod_{i\in\alpha} T^{a_i}; \alpha \in \mathcal{F}\}$. Without loss of generality we suppose $\mathrm{IP}-\lim_{T_\alpha\in\Sigma_\Gamma}\{\langle T_\alpha f, g\rangle - \langle f,1\rangle\langle 1,g\rangle\} = 0$ for all $f, g \in L^2(\mathbf{X},\mathcal{B},\mu)$ (cf. Theorem 1.2).

Let $\{\xi_k\}$ have the property described in Lemma 2.4. We define inductively a sequence $\{\alpha_n\}$ of pairwise disjoint finite subsets of $\mathbf{N}$ as follows. Let α_1 be any finite subset of $\mathbf{N}$ and suppose $\alpha_1,\ldots,\alpha_{n-1}$ have been defined. Let:

$$N_n = \max_{1\leq k\leq n} \#\Big\{\bigvee_{i=1}^{n-1} T_{\alpha_i}\xi_k\Big\}$$

and choose δ_n such that

$$|u-v| < \delta_n \text{ imples } |u\log u - v\log v| < \frac{1}{nN_n}.$$

Choose α_n such that $\alpha_n \cap \alpha_j = \emptyset$ for $1 \leq j \leq n-1$ and

$$|\mu(T_{\alpha_n}E\cap B) - \mu(E)\mu(B)| < \delta_n$$

for all $E \in \xi_k$, $B \in \bigvee_{i=0}^{n-1} T_{\alpha_i}\xi_k$ and $1 \leq k \leq n$. Then:

$$\Big|-\sum_{E,B}\mu(T_{\alpha_n}E\cap B)\ln\mu(T_{\alpha_n}E\cap B) + \sum_{E,B}\mu(E)\mu(B)\ln\mu(E)\mu(B)\Big| \leq \sum_{E,B}\frac{1}{nN_n}$$
$$\leq \sum_{E}\frac{1}{n}$$

https://doi.org/10.4153/CJM-1992-014-5 Published online by Cambridge University Press

where the sums are taken over all sets $E \in \xi_k$, $B \in \bigvee_{i=1}^{n-1} T^{t_i}\xi_k$. Note that:

$$\sum_{E,B} \mu(E)\mu(B)\ln\mu(E)\mu(B) = \sum_E \mu(E)\ln\mu(E) + \sum_B \mu(B)\ln\mu(B).$$

Therefore,

$$\lim_{n\to\infty}\Big\{-\sum_{E,B}\mu(T_{\alpha_n}E\cap B)\ln\mu(T_{\alpha_n}E\cap B) + \sum_E \mu(E)\ln\mu(E) + \sum_B \mu(B)\ln\mu(B)\Big\} = 0$$

i.e.

$$\lim_{n\to\infty}\Big\{H\Big(\bigvee_{i=1}^{n} T_{\alpha_i}\xi_k\Big) - H\Big(\bigvee_{i=1}^{n-1} T_{\alpha_i}\xi_k\Big)\Big\} = H(\xi_k)$$

for all ξ_k. Taking $s_n = \sum_{i\in\alpha_n} a_i$, we have $h_{\{s_n\}}(T,\xi_k) = H(\xi_k)$ for all k. For any finite partition ξ, by Lemma 2.4, there is a ξ_k such that $H(\xi \mid \xi_k) + H(\xi_k \mid \xi) < \varepsilon$. Since:

$$H\Big(\bigvee_{i=1}^{n} T^{s_i}(\xi\vee\xi_k)\Big) = H\Big(\bigvee_{i=1}^{n} T^{s_i}\xi\Big) + H\Big(\bigvee_{i=1}^{n} T^{s_i}\xi_k \Big| \bigvee_{i=1}^{n} T^{s_i}\xi\Big)$$

and

$$H\Big(\bigvee_{i=1}^{n} T^{s_i}(\xi\vee\xi_k)\Big) = H\Big(\bigvee_{i=1}^{n} T^{s_i}\xi_k\Big) + H\Big(\bigvee_{i=1}^{n} T^{s_i}\xi \Big| \bigvee_{i=1}^{n} T^{s_i}\xi_k\Big)$$

we have

$$\begin{aligned}\frac{1}{n}\Big|H\Big(\bigvee_{i=1}^{n} T^{s_i}\xi_k\Big) - H\Big(\bigvee_{i=1}^{n} T^{s_i}\xi\Big)\Big| &\le \frac{1}{n}\Big\{H\Big(\bigvee_{i=1}^{n} T^{s_i}\xi_k \,\Big|\, \bigvee_{i=1}^{n} T^{s_i}\xi\Big) + H\Big(\bigvee_{i=1}^{n} T^{s_i}\xi \,\Big|\, \bigvee_{i=1}^{n} T^{s_i}\xi_k\Big)\Big\}\\ &\le \frac{1}{n}\sum_{i=1}^{n}\Big(H(T^{s_i}\xi_k \mid T^{s_i}\xi) + H(T^{s_i}\xi \mid T^{s_i}\xi_k)\Big) < \varepsilon.\end{aligned}$$

Hence

$$|h_{\{s_n\}}(T,\xi) - h_{\{s_n\}}(T,\xi_k)| \le \varepsilon$$

which implies $|h_{\{s_n\}}(T,\xi) - H(\xi)| \le 2\varepsilon$ for all $\varepsilon > 0$, i.e. $h_{\{s_n\}}(T,\xi) = H(\xi)$.

Conversely if T is not mildly mixing, there is a nonconstant function f and a subset $\Gamma = \{t_n\}$ of $\mathbf{N}$ such that $\mathrm{IP}-\lim_{T_\alpha\in\Sigma_\Gamma} T_\alpha f = f$. By Proposition 1.3 there exists a sub-σ-algebra $\mathcal{B}_0$ such that for $f \in L^2(\mathbf{X},\mathcal{B}_0,\mu)$, $\mathrm{IP}-\lim_{T_\alpha\in\Sigma_\Gamma} T_\alpha f = f$. So $(\mathbf{X},\mathcal{B}_0,\mu,T)$ is rigid. By Theorem 0.3 for any $\mathcal{B}_0$ -measurable partition ξ, any sequence $\{\alpha_n\}$ of pairwise disjoint finite subsets of $\mathbf{N}$ and $s_n = \sum_{i\in\alpha_n} t_i$, $h_{\{s_n\}}(T,\xi) = 0$. But $\mathcal{B}_0$ is non-trivial. So there exists $\xi = \{A, A^c\}$ with $0 < \mu(A) < 1$. Hence $H(\xi) > 0$. This is a contradiction which implies our theorem.

Appendix.

SALESKI'S THEOREM. T is weakly mixing if and only if for any finite partition ξ, $\sup_{\Gamma\subset\mathbf{N}} h_\Gamma(T,\xi) = H(\xi)$.

A proof of Saleski's Theorem can be found in [8], p. 63.

PROOF OF THEOREM 0.2. Suppose T is weakly mixing. Then there is a subset $\mathbf{J}$ of $\mathbf{N}$ having density zero such that for any $A, B \in \mathcal{B}$ we have:

$$\lim_{\substack{n\to\infty \\ n\notin J}} \mu(T^{-n}A \cap B) = \mu(A)\mu(B).$$

Hence $\Gamma_0 = \Gamma \cap \mathbf{J}^c$ is an infinite subset of Γ. Let $\Gamma_0 = \{s_m\}$, then for all $A, B \in \mathcal{B}$ we have $\lim_{m\to\infty} \mu(T^{-s_m}A \cap B) = \mu(A)\mu(B)$. Taking a sequence $\{\xi_k\}$ of partitions as described in Lemma 2.5, we define a subset $\Gamma_1 = \{t_n\}$ of Γ_0 as follows. Let $t_1 = s_1$ and suppose that $t_1, t_2, \dots, t_{n-1}$ have been defined. Let

$$N_n = \max_{1\le k\le n} \#\Big\{\bigvee_{i=1}^{n-1} T^{-t_i}\xi_k\Big\}$$

and choose δ_n such that

$$|u - v| < \delta_n \text{ implies } |u\log u - v\log v| < \frac{1}{nN_n}.$$

Now choose t_n such that $t_n > t_{n-1}$ and

$$|\mu(T^{-t_n}A \cap B) - \mu(A)\mu(B)| < \delta_n$$

for all $A \in \xi_k$, $B \in \bigvee_{i=1}^{n-1} T^{-t_i}\xi_k$ and $1 \le k \le n$. Then:

$$\begin{aligned}\Big|-\sum_{A,B}\mu(T^{-t_n}A \cap B)\log\mu(T^{-t_n}A \cap B) + \sum_{A,B}\mu(A)\mu(B)\log\mu(A)\mu(B)\Big| &\le \sum_{A,B}\frac{1}{nN_n}\\ &\le \#\{\xi_k\}\frac{1}{n}\end{aligned}$$

where the sums are taken over all sets $A \in \xi_k$, $B \in \bigvee_{i=1}^{n-1} T^{-t_i}\xi_k$. Note that:

$$\begin{aligned}\sum_{A,B}\mu(B)\mu(B)\log\mu(A)\mu(B) &= \sum_{A,B}\mu(A)\mu(B)\log\mu(A) + \sum_{A,B}\mu(A)\mu(B)\log\mu(B)\\ &= \sum_{A}\mu(A)\log\mu(A) + \sum_{B}\mu(B)\log\mu(B).\end{aligned}$$

Therefore, it follows that

$$\lim_{n\to\infty}\Big\{H\Big(\bigvee_{i=1}^{n} T^{-t_n}\xi_k\Big) - H(\xi_k) - H\Big(\bigvee_{i=1}^{n-1} T^{-t_n}\xi_k\Big)\Big\} = 0.$$

So for all k, we have:

$$\lim_{n\to\infty}\Big\{H\Big(\bigvee_{i=1}^{n} T^{-t_n}\xi_k\Big) - H\Big(\bigvee_{i=1}^{n-1} T^{-t_n}\xi_k\Big)\Big\} = H(\xi_k).$$

Therefore, $h_{\Gamma_1}(T, \xi_k) = H(\xi_k)$ for all k. By the construction of $\{\xi_k\}$, we have $h_{\Gamma_1}(T, \xi) = H(\xi)$ for any finite partition ξ.

The proof in the other direction is contained in Saleski's Theorem.

References

1. H. Furstenberg, IP-*systems in ergodic theory*, Contemporary Mathematics, Conference in Modern Analysis and Prob. **26**(1984), 131–148.
2. ———, *Recurrence in ergodic theory and combinatorial number theory*. Princeton University Press, Princeton, N.J., 1981.
3. H. Furstenberg and B. Weiss, *The finite multipliers of infinite ergodic transformations*. In: Structure of attractors in dynamical systems. Lecture Notes in Mathematics **688**, Springer-Verlag, New York, 1978, 127–133.
4. P. Hulse, *Sequence entropy and subsequence generators*, J. London Math. Soc. (2)**26**(1982), 441–450.
5. A. A. Kirillov, *Dynamical systems, factors and representations of groups*, Russian Math. Surveys (5)**22** (1967), 63–75.
6. A. G. Kushnirenko, *On metric invariants of entropy type*, Russian Math. Surveys (5)**22**(1967), 53–62.
7. V. A. Rokhlin, *Lectures on the entropy theory of measure-preserving transformations*, Russian Math. Surveys (5)**22**(1967), 1–52.
8. A. Saleski, *Sequence entropy and mixing*, J. Math. Anal. Appl. **60**(1977), 58–66.
9. P. Walters, *An introduction to ergodic theory*. Springer-Verlag, New York, 1982.

The Ohio State University
100 Mathematics Building
231 West 18th Avenue
Columbus, Ohio 43210-1174
USA