On Spreading Partitions and Entropy in Infinite Measure Spaces

by

Eugene M. Klimko

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series No. 153
April 1968

On Spreading Partitions and Entropy in Infinite Measure Spaces

by

Eugene M. Klimko

1. Introduction.

In [7], the notion of a spreading partition was introduced and under certain regularity conditions, the pointwise convergence of information ratios was proved in the presence of an infinite invariant measure.

In this note, we show that these regularity conditions are satisfied for any partition having finite entropy and containing the first return partition. We also use this notion of a spreading partition to give a definition of entropy for sigma finite measure spaces which is at most equal to the Krengel entropy [8]; frequently the two entropies coincide. Our definition of entropy yields the equivalence of certain conditions of Krengel. We are also able to compute the entropy of a direct product under certain conditions.

2. Preliminary notions.

Let $(\Omega, Ⓐ, \mu)$ be a sigma finite measure space on which is defined an ergodic conservative endomorphism τ. The following definitions were introduced in [7]. The *information* of a partition Ⓕ $= \{F_i\}$ of a set $F \subset \Omega$, $0 < \mu(F) < \infty$, is defined by

$$(2.1) \qquad I(Ⓕ) = -\sum_i 1_{F_i} \log \mu(F_i) \quad .$$

The *entropy* of Ⓕ is

$$(2.2) \qquad H(Ⓕ) = \int_\Omega I(Ⓕ) \quad .$$

The refinement of $\textcircled{F}$ and $\textcircled{G}$ is the partition $\textcircled{F} \vee \textcircled{G}$ of $F \cup G$ consisting of sets of the form

$$F_i \cap G_j \, , \; F^c \cap G_j \, , \; F_i \cap G^c \, . \tag{2.3}$$

This method of refinement extends to any finite number of partitions and is easily seen to be commutative and associative.

Given a partition $\textcircled{F}$ of F and a sigma-field $\textcircled{G}$ of subsets of G, the <u>conditional information</u> of $\textcircled{F}$ given $\textcircled{G}$, $J(\textcircled{F} | \textcircled{G})$ is

$$J(\textcircled{F}|\textcircled{G}) = \begin{cases} \text{(i)} \; -\sum_i 1_{F_i} \log \mu(F_i | \textcircled{G}) & \text{on } F \cap G \\ \text{(ii)} \; -\log \mu(F^c | \textcircled{G}) & \text{on } F^c \cap G \\ \text{(iii)} \; -\sum_i 1_{F_i} \log \mu(F_i \cap G^c) & \text{on } F \cap G^c \\ \text{(iv)} \quad 0 & \text{on } F^c \cap G^c \end{cases} \tag{2.4}$$

This definition coincides with the usual definition on G but not on G^c. In the sequel, we shall use the following notation:

$$(\textcircled{F}_n)_i^j = \bigvee_{k=i}^{j} \tau^{-k} \textcircled{F}_n \, , \tag{2.5}$$

when there are no subscripts, we also omit parenthesis. For ergodic conservative endomorphisms, the partition $\textcircled{F}_1^n$ of $\bigcup_{i=1}^{n} \tau^{-i} F$ "spreads" over the entire space Ω. We now state some properties of the conditional information defined in (2.4).

Lemma 2.1 Let $\textcircled{F}, \textcircled{F}^1, \textcircled{F}^2$ be partitions of F, F^1 and F^2. Then

$$(2.6)\quad \begin{cases} \text{(i)} & I(\textcircled{F}^1 \vee \textcircled{F}^2) = I(\textcircled{F}^1) + J(\textcircled{F}^2 | \textcircled{F}^1) \\ \text{(ii)} & J(\textcircled{F}^1 \vee \textcircled{F}^2 | \textcircled{F}) = J(\textcircled{F}^1 | \textcircled{F}^2 \vee \textcircled{F}) + J(\textcircled{F}^2 | \textcircled{F}). \\ \text{(iii)} & \int_G J(\textcircled{F} | \textcircled{F}^1) \le \int_G J(\textcircled{F} | \textcircled{F}^2) \text{ for } \mu(G) < \infty \quad G \in \textcircled{F}^1 \subset \textcircled{F}^2 \\ \text{(iv)} & \lim_{n \to \infty} J(\textcircled{F} | \textcircled{F}_1^n) \text{ exists a.e.} \end{cases}$$

Corollary 2.1. Let $\textcircled{F}^1$ and $\textcircled{F}^2$ be partitions of F. Then

$$(2.7)\qquad H(\textcircled{F}^1 \vee \textcircled{F}^2) \le H(\textcircled{F}^1) + H(\textcircled{F}^2) + \mu(F) \log \mu(F) .$$

Proof. Let $\textcircled{F}^*$ be the partition consisting of the single set F itself. Then (2.7) follows from (2.6) (i), (iii) and the relation

$$\int_F J(\textcircled{F}^2 | \textcircled{F}^1) \le \int_F J(\textcircled{F}^2 | \textcircled{F}^*) = H(\textcircled{F}^2) + \mu(F) \log \mu(F) .$$

3. Main Results.

In [7], a ratio version of the Shannon Breimon McMillan theorem was proved for partitions $\textcircled{F}$ of F and $\textcircled{G}$ of G which satisfy the regularity condition

$$(3.1)\qquad \int_\Omega \sup_{n \ge 1} J(\textcircled{F} | \textcircled{F}_1^n) < \infty .$$

We shall give a sufficient condition for (3.1) to hold. We first need a modified version of a result due to Chung [3] (see also [2], [4]), which we state here without proof.

Lemma 3.1. (Chung) For $n = 1,2,\ldots$, let $(F)_n$ be an increasing sequence of partitions of F_n. Let (F) be a partition of F. Then for any $C_n \in (F)_n$ with $\mu(C_n) < \infty$, $C_n \subset F$ and $H((F) \cap C_n) < \infty$. Then there are constants α and β independent of C_n such that

$$(3.2) \quad \int_{C_n} \sup_{m \geq n} J((F) \mid (F)_1^m) \leq \mu(C_n)(\alpha + \beta \log \mu(C_n)) + 2H((F) \cap C_n) .$$

For any set F, the first return partition $(E)_F$ consists of the following sets: $E_0 = \phi$ and for $n = 1,2,\ldots$

$$(3.3) \qquad E_n = (F \cap \tau^{-n} F) - \bigcup_{i=0}^{n-1} E_i \quad .$$

For ergodic conservative endomorphisms, the first return sets E_n have the properties

$$(3.4) \qquad F = \bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} \tau^{-n} E_n , \quad F^c = \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{n-1} \tau^{-j} E_n \quad .$$

Moreover, the sets in the double union are mutually disjoint (c.f. [6]) .

Theorem 3.1. Let $(F) = \{F_m\}_{m=1}^{\infty}$ be a partition of F. If $(E)_F \subset (F)$ and $H((F)) < \infty$, then (3.1) holds.

Proof. Let $f_n = J((F) \mid (F)_1^n)$. We have

$$(3.5) \qquad \int_{F^c} \sup_{n \geq 1} f_n = \sum_{j=2}^{\infty} \sum_{i=1}^{j-1} \int_{\tau^{-i} E_j} \sup_{n \geq 1} f_n \quad .$$

Fix i, j, $0 < i < j$, $j \geq 2$. For $n < i$, $\tau^{-i} E_j \subset \bigcap_{k=0}^{n} \tau^{-k} F^c$. Hence, $f_n = 0$ on $\tau^{-i} E_j$ by case (iv) of definition (2.4). On the other hand, if $n \geq i$, $\tau^{-i} E_j \in \mathcal{F}_1^n$, and $\tau^{-i} E_j \subset F^c$, which implies that $\mu(F^c | \mathcal{F}_1^n) = 1$ and $f_n = 0$ on $\tau^{-i} E_j$. Therefore, each of the integrals in (3.5) is 0. Let $E_j^2 = \tau^{-j} E_j$; we have

$$(3.6) \quad \int_\Omega \sup_{n\geq 1} f_n = \sum_{j=1}^{\infty} \int_{E_j^2} \sup_{n\geq 1} f_n \leq \sum_{j=1}^{\infty} \int_{E_j^2} \sup_{1\leq n\leq j-1} f_n + \sum_{j=1}^{\infty} \int_{E_j^2} \sup_{n\geq j} f_n \; .$$

Since each point in E_j "returns" to F for the first time under τ^{-j}, $E_j^2 \subset \bigcap_{k=1}^{n} \tau^{-k} F^c$ for $n < j$. On the set E_j^2, by definition (2.4)(iii) we have

$$(3.7) \qquad f_n = -\sum_{m=1}^{\infty} 1_{F_m \cap E_j^2} \log \mu(F_m \cap \bigcap_{k=1}^{n} \tau^{-k} F^c)$$

for $1 \leq n \leq j - 1$. Furthermore, on this set, $f_n \geq f_{n-1}$ for $2 \leq n \leq j - 1$. We have

$$(3.8) \quad \sup_{1\leq n\leq j-1} f_n = -\sum_{m=1}^{\infty} 1_{F_m \cap E_j^2} \log \mu(F_m \cap \bigcap_{k=1}^{j=1} \tau^{-k} F^c) \leq$$

$$-\sum_{m=1}^{\infty} 1_{F_m \cap E_j^2} \log \mu(F_m \cap E_j^2) \qquad .$$

Therefore, the first sum on the right of (3.6) is bounded by

$$(3.9) \qquad H(\textcircled{F} \vee \textcircled{E}_F^2) \leq H(\textcircled{F}) + H(\textcircled{E}_F) + \mu(F) \log \mu(F)$$

where $\textcircled{E}_F^2 = \{E_j^2\}$ is the second return partition which has the same entropy as the first return partition; the inequality follows from corollary 2.1. Since $E_j^2 \in \textcircled{F}_1^n$ for $n \geq j$, we have by Lemma 2.3

$$(3.10) \quad \int_{E_j^2} \sup_{n \geq j} f_n \leq \mu(E_j^2)(\alpha + \beta \log \mu(E_j^2)) + H(\textcircled{F} \cap E_j^2) .$$

We sum (3.10) on j and combine with (3.9) to obtain

$$(3.11) \quad \int_\Omega \sup_{n \geq 1} f_n \leq \alpha\mu(F) + 2H(\textcircled{F}) + (2-\beta)H(\textcircled{E}_F) + \mu(F) \log \mu(F) .$$

We now consider some properties of $\int_\Omega J(\textcircled{F} | \textcircled{F}_1^\infty)$.

<u>Lemma 3.2.</u> Let F be a set and let $\textcircled{F}$ be any partition of F. Then for $n = 1,2,\ldots,$ $\tau^{-n} \textcircled{E}_F \subset \textcircled{F}_1^\infty$; in particular, $F \in \textcircled{F}_1^\infty$.

<u>Proof</u>. $\tau^{-n} E_1 = \tau^{-n} F \cap \tau^{-n-1} F \in \textcircled{F}_1^\infty$. The first inclusion follows by induction on E_m, since

$$(3.12) \qquad \tau^{-n} E_m = (\tau^{-n} F \cap \tau^{-n-m} F) - \bigcup_{i=0}^{m-1} \tau^{-n} E_i \quad .$$

The second conclusion holds because $F = \bigcup_{n=1}^{\infty} \tau^{-n} E_n$.

By arguments similar to those used in the beginning of the proof of theorem 3.1, we obtain:

<u>Corollary 3.1</u>. $J(\textcircled{F} | \textcircled{F}_1^\infty) = 0$ on F^c.

Corollary 3.2. If $\textcircled{F}$ is any partition of F and $H(\textcircled{F})$, $H(\textcircled{E}_F)$ are both finite, then

$$(3.13) \qquad \int_\Omega J(\textcircled{E}_F \vee \textcircled{F} \mid (\textcircled{F} \vee \textcircled{E}_F)_1^\infty) = \int_\Omega J(\textcircled{F} \mid \textcircled{F}_1^\infty) < \infty .$$

We may now define the entropy of τ in the set F, $h^*(F,\tau)$ by

$$(3.14) \qquad h^*(F,\tau) = \sup_{\textcircled{F}} \int_F J(\textcircled{F} \mid \textcircled{F}_1^\infty)$$

where supremum is taken over all partitions of F which have finite entropy. The entropy of τ, $h^*(\tau)$ is

$$(3.15) \qquad h^*(\tau) = \sup_{\mu(F)<\infty} h^*(F,\tau) .$$

Let E_n be the first return sets of F. The endomorphism τ of Ω induces an endomorphism τ_F on F defined by

$$(3.16) \qquad \tau_F(w) = \tau^n(w) \quad \text{for } w \in E_n .$$

For sigma finite measure spaces, Krengel [8] defined the entropy of τ as

$$(3.17) \qquad h(\tau) = \sup_{\mu(F)<\infty} h(\tau_F)$$

where $h(\tau_F)$ is the entropy of the induced transformation τ_F on the finite measure space $(F, \textcircled{A} \cap F, \mu)$.

Theorem 3.2. $h^*(\tau) \leq h(\tau)$. Moreover, if there is a set whose first return partition has finite entropy, then $h^*(\tau) = h(\tau)$.

Proof. Let F be any set of finite measure and $Ⓕ$ a partition of F such that $H(Ⓕ) < \infty$. We consider

$$(3.18)\quad \bigvee_{i=1}^{\infty} \tau_F^{-i}\, Ⓕ \subset \bigvee_{i=1}^{\infty} \tau_F^{-i}\,(Ⓕ \vee Ⓔ_F) = \bigvee_{i=1}^{\infty} \tau^{-i}(Ⓕ \vee Ⓔ_F) = \bigvee_{i=1}^{\infty} \tau^{-i}\, Ⓕ .$$

The first inclusion is obvious and the last equality follows from Lemma 3.2. The relation

$$(3.19)\qquad F \cap \bigvee_{i=1}^{\infty} \tau^{-i}(Ⓕ \vee Ⓔ_F) \subset \bigvee_{i=1}^{\infty} \tau_F^{-i}(Ⓕ \vee Ⓔ_F)$$

has been used by Scheller [9, p. 44] (see also [8]). The reverse inclusion follows essentially from Lemma 5 of [1] where it is assumed that τ is an automorphism, however the same proof may be used for endomorphism's. From (3.18) and Lemma 2.1 (iii), it follows that

$$(3.20)\qquad \int_F J(Ⓕ \mid Ⓕ_1^{\infty}) \leq \int_F J(Ⓕ \mid \bigvee_{i=1}^{\infty} \tau_F^{-i}\, Ⓕ) .$$

We note that the expression to the right in (3.20) is the usual conditional entropy of the partition $Ⓕ$ of the finite measure space $(F, F \cap Ⓐ, \mu)$. Taking supremum over all partitions $Ⓕ$ of F , $H(Ⓕ) < \infty$, we obtain

$$(3.21)\qquad h(F,\tau) \leq h(\tau_F) \leq h(\tau) .$$

Therefore, $h^*(\tau) \leq h(\tau)$. To prove the second assertion, let F be such that $H(Ⓔ_F) < \infty$. To obtain the entropy $h(\tau_F)$ when $H(Ⓔ_F) < \infty$, it suffices to take supremum over all partitions of the type $Ⓕ \vee Ⓔ_F$, $H(Ⓕ) < \infty$. For these partitions, the inclusion in (3.18) becomes equality as well as the inequality in (3.20) and we obtain

$$(3.22) \qquad h(F,\tau) = h(\tau_F) \; .$$

Krengel has shown ([8] Lemma 3.1) that for ergodic conservative endomorphisms $h(\tau) = h(\tau_F)$ for any set F for which $0 < \mu(F) < \infty$. Therefore, by (3.22) $h^*(\tau) \geq h(\tau)$ which completes the proof.

<u>Remark</u>. Krengel does not assume that τ is ergodic. Instead, he considers sweep out sets; that is, sets F for which $\bigcup_{i=0}^{\infty} \tau^{-i} F = \Omega$. For nonergodic transformations, the first assertion of theorem 3.2 remains valid and if there is a <u>sweep out</u> set whose first return partition has finite entropy, then $h^*(\tau) = h(\tau)$.

Consider the following conditions:

(I) $\quad h(\tau) = 0$

(I') $\quad h^*(\tau) = 0$

(II) For every finite partition $Ⓕ^*$ of Ω of the form

$$Ⓕ^* = \{F_1, \ldots F_n, \Omega - \bigcup_{j=1}^{n} F_j\} \text{ with } \mu(F_i) < \infty \quad i = 1,\ldots,n ,$$

$$(3.23) \qquad Ⓕ^* \subset \bigvee_{i=1}^{\infty} \tau^{-i} Ⓕ^*$$

(III) (3.22) holds for every countable partition Ⓕ of Ω of the form

$$Ⓕ = \{F_1, F_2, \ldots, \Omega - \bigcup_{j=1}^{\infty} F_j\} \text{ with } \mu(\bigcup_{j=1}^{\infty} F_j) < \infty \text{ and } H(Ⓕ \cap \bigcup_{j=1}^{\infty} F_j) < \infty \; .$$

Krengel showed that (I) implies (II) and (III) and that if there is a set F whose first return partition has finite entropy, then (I), (II) and (III) are equivalent.

<u>Theorem 3.3</u>. Conditions (I'), (II) and (III) are equivalent.

<u>Proof</u>. (I') $\Rightarrow$ (II). The sets F_i, $i = 1, 2, \ldots, n$ form a finite partition $\textcircled{F}$ of $\bigcup_{i=1}^{n} F_i$. Since $h^*(\tau) = 0$, $J(\textcircled{F} \mid \textcircled{F}_1^\infty) = 0$ which implies that $\textcircled{F} \in \textcircled{F}_1^\infty$. Let $\textcircled{F}^*$ be the partition in (II). Since τ is ergodic and conservative, $\bigcup_{i=1}^{\infty} \tau^{-i}\,\textcircled{F} = \Omega$ and $\textcircled{F}_1^\infty = (\textcircled{F}^*)_1^\infty$.

(II) $\Rightarrow$ (III) is trivial.

(III) $\Rightarrow$ (I). Let F_i $i = 1, 2, \ldots$ be such that $-\Sigma\,\mu(F_i)\log\mu(F_i) < \infty$. The F_i form a partition $\textcircled{F}$ of $\cup F_i$ for which $\textcircled{F} \in \textcircled{F}_1^{*\infty} = \textcircled{F}_1^\infty$ which implies that $J(\textcircled{F} \mid \textcircled{F}_1^\infty) = 0$. Since this is true for every partition of $\cup F_i$ with finite entropy, $h(\cup F_i, \tau) = 0$. $h(F,\tau) = 0$ for every set of finite measure, hence $h^*(\tau) = 0$. In general, the question of computing $h^*(\tau)$ when $h^*(\tau) \neq h(\tau)$ is open as well as the question of obtaining a maximal invariant sigma field $\textcircled{B}_N$ for which $h^*(\tau) = 0$.

4. <u>Entropy of a direct product</u>.

We compute the entropy of a direct product of two endomorphisms τ_X and τ_Y of the measure spaces $(X, \textcircled{X}, \mu)$ and $(Y, \textcircled{Y}, \lambda)$ when $h^*(\tau_Y) = 0$ and $\lambda(Y) < \infty$. Examples of Kakutani and Parry show that if $\mu(X) = \lambda(Y) = \infty$, then $\tau_X \times \tau_Y$ need not be ergodic.

<u>Theorem 4.1</u>. Let $\tau_X \times \tau_Y$ be an ergodic endomorphism of the product $(X \times Y, \textcircled{X} \times \textcircled{Y}, \mu \times \lambda)$. Let $F \subset X$ be a set whose first return partition

has finite entropy; let $\lambda(Y) < \infty$ and $h^*(\tau_Y) = 0$. Then

$$(4.1) \qquad h^*(\tau_X) = \lambda(Y)\, h^*(\tau_X) \; .$$

Proof. Let E_n $n = 1,2,\ldots$ be the first return partition of F . Then, the first return partition of $F \times Y$ consists of the sets: $E_n \times Y$, $n = 1,2,\ldots$. To compute the entropy of $\tau_{F \times Y}$, the transformation induced on $F \times Y$ by $\tau_X \times \tau_Y$, it suffices to consider partitions of $F \times Y$ consisting of rectangles (see [5], p. 277 f.f.). Moreover, it suffices to consider partitions which contain the first return partitions of $F \times Y$. We have equality in (3.20) and therefore

$$(4.2) \qquad h^*(\tau_X \times \tau_Y) = \sup_{F \times G} \int_{F \times Y} J(F \times G \,|\, (F \times G)_1^\infty) \; .$$

where $F = \{F_i\}$ partitions F, $E_n \in F$, $n = 1,2,\ldots$ and $G = \{G_i\}$ partitions Y . The partitions

$$(4.3) \qquad \bigvee_{i=1}^{n} (\tau_X \times \tau_Y)^{-i} \, (F \times G) = (\bigvee_{i=1}^{n} \tau_X^{-i} F) \times (\bigvee_{i=1}^{n} \tau_Y^{-i} G)$$

generate $\bigvee_{i=1}^{\infty} (\tau_X \times \tau_Y)^{-i} \, (F \times G)$ and the factors on the right are independent sigma fields with respect to the measure $\mu \times \lambda$. Therefore, the conditional measures multiply

$$(4.4) \quad \mu \times \lambda(F_j \times G_k \,|\, \bigvee_{i=1}^{\infty} (\tau_X \times \tau_Y)^{-i} (F \times G)) = \mu(F_j \,|\, F_1^\infty) \cdot \lambda(G_k \,|\, G_1^\infty)$$

and we obtain

$$(4.5) \quad \int_{F \times Y} J\left(\textcircled{F} \times \textcircled{G} \,\Big|\, \bigvee_{i=1}^{\infty} (\tau_X \times \tau_Y)^{-i} (\textcircled{F} \times \textcircled{G})\right) =$$

$$\lambda(Y) \int_F J(\textcircled{F} \,|\, \textcircled{F}_1^{\infty}) + \mu(F) \int_Y J(\textcircled{G} \,|\, \textcircled{G}_1^{\infty}) \quad .$$

The last term in (4.5) is zero; we take supremum over rectangular partitions of $F \times Y$ to obtain

$$h(\tau_X \times \tau_Y) = \lambda(Y)\, h(\tau_X) \quad .$$

The proof of the preceeding theorem indicates that in general, the entropy of a direct product of two infinite measure spaces is infinite whenever the product transformation is ergodic.

5. Concluding remarks.

The case when there is no set F whose first return partition has finite entropy remains to be investigated. An approximation result is needed to show that one need only consider partitions which are dense in a given sigma field. Such a result would permit the characterization of entropy zero in terms of a maximal sigma field $\textcircled{B}_N$. The basic difficulty is that known results rely on the L_1 version of the Shannon Breiman McMillan theorem.

References

1) Abramov, L. M.: The entropy of a derived automorphism, Doklady Akad. Nauk SSSR 128 647-650, (1959), (Russian). (Am. Math. Soc. Translations, ser 2, **49** 162-166, (1966).)

2) Breiman, L.: The individual ergodic theorem of information theory. Ann. Math. Statist. **28**, 809-811, (1957); correction note, ibid. **31**, 809-810, (1960).

3) Chung, K.L.: A note on the ergodic theorem of information theory. Ann. Math. Statist. **32**, 612-614 (1961).

4) Ionescu Tulcea, A.: Contributions to information theory for abstract alphabets. Arkiv för Matematik **4**, 235-247 (1961).

5) Jacobs, K.: Lecture notes on ergodic theory. Universitet Aarhus, Matematisk Institut, (1962/63).

6) Kakutani, S.: Induced measure preserving transformations. Proc. Imp. Acad. Sci. Tokyo **19** 635-641 (1943).

7) Klimko, E.M. and Sucheston, L.: On convergence of information in spaces with infinite invariant measure. To appear: Z. Wahrscheinlichkeitstheorie verw. Geb.

8) Krengel, U.: Entropy of conservative transformations. Z. Wahrscheinlichkeitstheorie verw. Geb **7**, 161-181 (1967).

9) Scheller, H.: Induzierte dynamische Systeme, Diplomarbeit der Mathematisch-Naturwissenschaftlichen Fakultat der Universitat Goettingen, 62 pages, (1965).