A NOTE ON THE ENTROPY OF SKEW PRODUCT TRANSFORMATIONS[1]

ROY L. ADLER

Let $(X, \mathfrak{X}, \lambda)$ and $(Y, \mathcal{Y}, \mu)$ be two Lebesgue spaces with $\mathfrak{X}$ and $\mathcal{Y}$ the fields of measurable subsets of X and Y respectively. λ and μ are countably additive measures on $\mathfrak{X}$ and $\mathcal{Y}$ respectively with $\lambda(X) = \mu(Y) = 1$. Let $(Z, \mathcal{Z}, \nu) = (X \times Y, \mathfrak{X} \times \mathcal{Y}, \lambda \times \mu)$ denote the direct product of the above measure spaces. Let ϕ be a measure preserving transformation on X, and for each $x \in X$ let ψ_x be measure preserving transformation on Y. If the family $\{\psi_x : x \in X\}$ of measure preserving transformations satisfies certain measurability conditions (see [2, pp. 83, 84]), then it can be shown that the transformation T defined by

$$T(x, y) = (\phi x, \psi_x y)$$

is a measure preserving transformation on Z. T is called the skew product transformation of ϕ with the family $\{\psi_x : x \in X\}$.

The purpose of this work is to compute the entropy $h(T)$. (For definition of entropy of a measure preserving transformation and the associated notation consult [3] and [4].) The natural conjecture is

$$(*) \qquad h(T) = h(\phi) + \int_X h(\psi_x)\lambda(dx).$$

This conjecture is substantiated in several instances. When $\psi_x = \psi$ for all $x \in X$, (*) reduces to the formula for direct product transformations (see [4] formula (β)); i.e.,

$$h(T) = h(\phi) + h(\psi).$$

For $\phi = I$ the identity transformation on X, (*) reduces to the case of decomposition of a measure preserving transformation into components (see [4] formula (ϵ)); i.e.,

$$h(T) = \int_X h(\psi_x)\lambda(dx).$$

When Y = unit interval and $\psi_x y = y + \alpha(x)$ (mod 1) where $\alpha(\cdot)$ is some real-valued measurable function on X, Abramov [1] has shown

$$h(T) = h(\phi)$$

Received by the editors May 2, 1962.

[1] The material in this paper was included in the author's doctoral dissertation submitted to Yale University (1961) under the direction of Professor S. Kakutani.

which is again a special case of (*) since $h(\psi_x)=0$, $x\in X$.

In general (*) is not true. However, we shall derive a formula which differs from (*) in the function occurring within the integral.

Let $\mathfrak{X}_k$, $k=1, 2, \cdots$ be an increasing sequence of finite subfields of $\mathfrak{X}$ whose union generates $\mathfrak{X}$ and let $\mathfrak{Z}_m$, $m=1, 2, \cdots$ be an increasing sequence of finite subfields of $\mathfrak{Z}$ whose union generates $\mathfrak{Z}$. Let ${}_T^n\mathfrak{Z}_m$ denote $\mathfrak{Z}_m \vee T\mathfrak{Z}_m \vee \cdots \vee T^{n-1}\mathfrak{Z}_m$. Denote by $(\mathfrak{Z}_m)_x$ the field of subsets of Y which consists of x-sections of sets in $\mathfrak{Z}_m$. We observe

$$({}_T^n\mathfrak{Z}_m)_x = (\mathfrak{Z}_m)_x \vee \psi_{\phi^{-1}x}(\mathfrak{Z}_m)_{\phi^{-1}x} \vee \cdots \vee \psi_{\phi^{-1}x}\psi_{\phi^{-2}x} \cdots \psi_{\phi^{-n+1}x}(\mathfrak{Z}_m)_{\phi^{-n+1}x}. \tag{1}$$

We shall employ an ambiguity whose meaning will be clear in context by having the symbols $\mathfrak{X}_k$ represent either fields of measurable subsets of X or fields of cylinder sets in $\mathfrak{Z}$ based on subsets of X in $\mathfrak{X}_k$. Keeping this ambiguity in mind consider the following relation between mean entropy and mean (conditional) entropy of finite fields (see [4, p. 980])

$$\begin{aligned} H({}_T^n\mathfrak{X}_k \vee {}_T^n\mathfrak{Z}_m) &= H({}_T^n\mathfrak{X}_k) + H({}_T^n\mathfrak{Z}_m \mid {}_T^n\mathfrak{X}_k) \\ &= H({}_\phi^n\mathfrak{X}_k) + H({}_T^n\mathfrak{Z}_m \mid {}_T^n\mathfrak{X}_k). \end{aligned}$$

Dividing by n and observing ${}_T^n\mathfrak{X}_k \subseteq \mathfrak{X}$ we have the inequality

$$\frac{H({}_T^n\mathfrak{X}_k \vee {}_T^n\mathfrak{Z}_m)}{n} \geqq \frac{H({}_\phi^n\mathfrak{X}_k)}{n} + \frac{H({}_T^n\mathfrak{Z}_m \mid \mathfrak{X})}{n}. \tag{2}$$

The definition of mean (conditional) entropy yields

$$H({}_T^n\mathfrak{Z}_m \mid \mathfrak{X}) = \int_X H(({}_T^n\mathfrak{Z}_m)_x)\lambda(dx).$$

By replacing the function in the integral with (1) and substituting in (2) we get

$$\begin{aligned} \frac{H({}_T^n\mathfrak{X}_k \vee {}_T^n\mathfrak{Z}_m)}{n} &\geqq \frac{H({}_\phi^n\mathfrak{X}_k)}{n} \\ &+ \int_X \frac{H((\mathfrak{Z}_m)_x \vee \psi_{\phi^{-1}x}(\mathfrak{Z}_m)_{\phi^{1-x}} \vee \cdots \vee \psi_{\phi^{-1}x}\psi_{\phi^{-2}x} \cdots \psi_{\phi^{-n+1}x}(\mathfrak{Z}_m)_{\phi^{-n+1}x})}{n} \\ &\cdot \lambda(dx). \end{aligned} \tag{3}$$

The following identity can be established (see for instance [3, p. 33]):

$$H((\mathcal{Z}_m)_x \vee \psi_{\phi^{-1}x}(\mathcal{Z}_m)_{\phi^{-1}x} \vee \cdots \vee \psi_{\phi^{-1}x}\psi_{\phi^{-2}x} \cdots \psi_{\phi^{-n+1}x}(\mathcal{Z}_m)_{\phi^{-n+1}x})$$
$$= \sum_{j=0}^{n-1} H((\mathcal{Z}_m)_{\phi^{-j}x} \mid \psi_{\phi^{-j-1}x}(\mathcal{Z}_m)_{\phi^{-j-1}x} \vee \cdots \vee \psi_{\phi^{-j-1}x} \cdots \psi_{\phi^{-n+1}x}(\mathcal{Z}_m)_{\phi^{-n+1}x}). \tag{4}$$

As n tends to ∞,

$$H((\mathcal{Z}_m)_{\phi^{-j}x} \mid \psi_{\phi^{-j-1}x}(\mathcal{Z}_m)_{\phi^{-j-1}x} \vee \cdots \vee \psi_{\phi^{-j-1}x} \cdots \psi_{\phi^{-n+1}x}(\mathcal{Z}_m)_{\phi^{-n+1}x})$$

decreases to a limit which we denote by $f_\phi(x, \mathcal{Z}_m, j)$. Likewise, as $n \to \infty$,

$$\int_X H((\mathcal{Z}_m)_{\phi^{-j}x} \mid \psi_{\phi^{-j-1}x}(\mathcal{Z}_m)_{\phi^{-j-1}x} \vee \cdots \vee \psi_{\phi^{-j-1}x} \cdots \psi_{\phi^{-n+1}x}(\mathcal{Z}_m)_{\phi^{-n+1}x})\lambda(dx)$$

tends to

$$\int_X f_\phi(x, \mathcal{Z}_m, j)\lambda(dx).$$

By virtue of the fact that ϕ is measure preserving it follows that

$$\int_X f_\phi(x, \mathcal{Z}_m, j)\lambda(dx) = \int_X f_\phi(x, \mathcal{Z}_m, 0)\lambda(dx). \tag{5}$$

Having outgrown the need for a function of three variables we replace $f_\phi(x, \mathcal{Z}_m, 0)$ by simply $f_\phi(x, \mathcal{Z}_m)$. From (4) and (5) and the fact that ordinary convergence implies Cesaro convergence we obtain

$$\int_X \frac{H((\overset{n}{{}_T}\mathcal{Z}_m)_x)}{n}\lambda(dx) \to \int_X f_\phi(x, \mathcal{Z}_m)\lambda(dx) \tag{6}$$

as $n \to \infty$. Taking limits in (3) as $n \to \infty$ yields

$$h(T) \geqq h(T, \mathfrak{X}_k \vee \mathcal{Z}_m) \geqq h(\phi, \mathfrak{X}_k) + \int_X f_\phi(x, \mathcal{Z}_m)\lambda(dx). \tag{7}$$

Then letting $k \to \infty$ we get

$$h(T) \geqq h(\phi) + \int_X f(x, \mathcal{Z}_m)\lambda(dx). \tag{8}$$

In order to obtain the reverse inequality consider

$$
(9)\qquad
\begin{aligned}
\frac{H({}_T^{nl}\mathcal{Z}_m)}{nl} &\leqq \frac{H({}_T^{nl}\mathfrak{X}_k \vee {}_T^{nl}\mathcal{Z}_m)}{nl} = \frac{H({}_T^{nl}\mathfrak{X}_k)}{nl} + \frac{H({}_T^{nl}\mathcal{Z}_m \mid {}_T^{nl}\mathfrak{X}_k)}{nl} \\
&\leqq \frac{H({}_\phi^{nl}\mathfrak{X}_x)}{nl} + \frac{H({}_T^{nl}\mathcal{Z}_m \mid {}_T^{nl}\mathfrak{X}_k)}{nl}.
\end{aligned}
$$

Now

$$
(10)\qquad
\begin{aligned}
H({}_T^{nl}\mathcal{Z}_m \mid {}_T^{nl}\mathfrak{X}_k) &\leqq \sum_{i=0}^{n-1} H(T^{il}({}_T^{l}\mathcal{Z}_m) \mid {}_T^{nl}\mathfrak{X}_k) \leqq \sum_{i=0}^{n-1} H(T^{il}({}_T^{l}\mathcal{Z}_m) \mid T^{il}\mathfrak{X}_k) \\
&\leqq nH({}_T^{l}\mathcal{Z}_m \mid \mathfrak{X}_k).
\end{aligned}
$$

Combining (9) and (10) we have

$$
(11)\qquad \frac{(H{}_T^{nl}\mathcal{Z}_m)}{nl} \leqq \frac{H({}_\phi^{nl}\mathfrak{X}_k)}{nl} + \frac{H({}_T^{l}\mathcal{Z}_m \mid \mathfrak{X}_k)}{l}.
$$

Letting $n \to \infty$ in (11)

$$
(12)\qquad h(T, \mathcal{Z}_m) \leqq h(\phi, \mathfrak{X}_k) + \frac{H({}_T^{l}\mathcal{Z}_m \mid \mathfrak{X}_k)}{l},
$$

and letting $k \to \infty$ in (12)

$$
h(T, \mathcal{Z}_m) \leqq h(\phi) + \frac{H({}_T^{l}\mathcal{Z}_m \mid \mathfrak{X})}{l},
$$

or equivalently,

$$
(13)\qquad h(T, \mathcal{Z}_m) \leqq h(\phi) + \int_X \frac{H(({}_T^{l}\mathcal{Z}_m)_x)}{l}\,\lambda(dx).
$$

Next letting $l \to \infty$ and combining the result with (8) we have

$$
(14)\qquad h(T, \mathcal{Z}_m) \leqq h(\phi) + \int_X f_\phi(x, \mathcal{Z}_m)\lambda(dx) \leqq h(T).
$$

Now

$$
f_\phi(x, \mathcal{Z}_m) = \lim_{n \to \infty} \frac{H(({}_T^{n}\mathcal{Z}_m)_x)}{n},
$$

and it is clear that $f(x, \mathcal{Z}_m)$ increases with m to a possibly infinite but well defined limit $f_\phi(x)$. Since $\lim_{m\to\infty} h(T, Z_m) = h(T)$, it follows from (14) that

$$(**) \qquad h(T) = h(\phi) + \int_X f_\phi(x)\lambda(dx).$$

Of course we would like to establish

$$(***) \qquad \int_X f_\phi(x)\lambda(dx) = \int_X h(\psi_x)\lambda(dx)$$

where

$$f_\phi(x) = \lim_{m\to\infty} f_\phi(x, Z_m)$$

$$f_\phi(x, Z_m) = \lim_{n\to\infty} \frac{H((Z_m)_x \vee \psi_{\phi^{-1}x}(Z_m)_{\phi^{-1}x} \vee \cdots \vee \psi_{\phi^{-1}x}\psi_{\phi^{-2}x}\cdots \psi_{\phi^{-n+1}x}(Z_m)_{\phi^{-n+1}x})}{n}$$

$$h(\psi_x) = \lim_{m\to\infty} h(\psi_x, Z_m)$$

$$h(\psi_x, Z_m) = \lim_{n\to\infty} \frac{H((Z_m)_x \vee \psi_x(Z_m)_x \vee \cdots \vee \psi_x^{n-1}(Z_m)_x)}{n}.$$

The quantities $h(\psi_x, Z_m)$ and $f(x, Z_m)$ are different by the nature of their definitions. Perhaps only mild restrictions are required so that the differences can be eliminated by integration to yield (***). The following example, however, reveals that in general (***) is not true: let $X = X_1 \cup X_2$ where $m(X_1) = m(X_2) = \frac{1}{2}$; let $\psi_x = \psi$, $x \in X_1$ and $\psi_x = \psi^{-1}$, $x \in X_2$ where ψ is a measure preserving transformation on Y such that $h(\psi) \neq 0$; and let ϕ be a measure preserving transformation on X such that $\phi X_1 = X_2$, $\phi X_2 = X_1$ and $\phi^2 = I$. Then for T: $(x, y) \to (\phi x, \psi_x y)$ we have T^2 is the identity transformation on $X \times Y$ so that $h(T) = 0$; but $h(\phi) + \int_X h(\psi_x)\lambda(dx) = h(\psi) \neq 0$.

References

1. L. M. Abramov, *On entropy of flows*, Dokl. Akad. Nauk SSSR **128** (1959), 873–875.

2. H. Anzai, *Ergodic skew product transformations on the torus*, Osaka Math. J. **3** (1959), 83–99.

3. P. R. Halmos, *Entropy in ergodic theory*, Univ. of Chicago, 1959. (Mimeographed Notes)

4. V. A. Rokhlin, *Entropy of metric automorphism*, Dokl. Akad. Nauk SSSR **124** (1959), 768–771.

International Business Machines Corporation