J. Korean Math. Soc. **32** (1995), No. 3, pp. 573–582

CONTINUITY OF DIRECTIONAL ENTROPY FOR A CLASS OF Z^2–ACTIONS*

KYEWON K. PARK

1. Introduction

J.Milnor[Mi2] has introduced the notion of directional entropy in his study of Cellular Automata. Cellular Automaton map can be considered as a continuous map from a space K^{Z^n} to itself which commute with the translation of the lattice Z^n. Since the space K^{Z^n} is compact, map S is uniformly continuous. Hence S is a block map(a finite code)[He]. (S is said to have a finite memory.) In the case of $n = 1$, we have a shift map, T on K^Z, and a block map S and they together generate a Z^2 action.

Given an arbitrary Z^2 action, it is not hard to define the entropy of a Z-action T^kS^ℓ, where k and ℓ are both integers. Since the directional entropy is defined in all directions besides the rational directions like $(k, \ell) \in Z^2$, it can be considered as a generalization of the entropy of non cocompact subgroups. It is shown in [Si] and [Pal] that the directional entropy function is upper semi-continuous in the case of a Z^2 action generated by a Cellular Automaton map like Milnor's setting. It is also shown by Thouvenot that the directional entropy is not upper semi-continuous in a more general setting.

Also the notion of directional entropy is an important tool for the computation of the entropy of a skew product with a more general group action. That is, if we have a Z^2 action generated by a Cellular Automaton map on the fiber Y and a skewing function $\varphi : X \to Z^2$, the entropy of the skew product on $X \times Y$ is computable in terms of the entropy of the base transfomation and the integral of the skewing

Received July 6, 1994.
1991 AMS Subject Classification: Primary 28D05.
Key words: directional entropy, cone entropy.
This work was partially supported by Ajou University Faculty Grant and GARC-KOSEF.

function. We require the skewing function to be integrable.(See[Pa2] and [Pa3] for details.)

Given an ergodic system of finite entropy $(X, T, \mathcal{F}, \mu)$, it can be represented (*isomorphic*) as a shift space of finite alphabets. Without confusion, we will denote this symbolic system by $(X, T, \mathcal{F}, \mu)$. Let S be a measure preserving invertable map of X generated by a block map. Hence $\{T, S\}$ generate a Z^2-action on X. By a block map S, we mean a map determined by a given rule f. That is, the n^{th} coordinate of Sx is determined by $f(x_{n-s}, x_{n-s+1}, \cdots, x_0, \cdots, x_{n+s})$. We call S a block map of size $2s$.

In this paper we will show that the upper semi-continuity holds for a more general class of Z^2 actions than the class shown in [Pa1]. In particular, if S is a finitary code with finite expected code length, then the directional entropy is upper semi -continuous. It is not yet clear if the directional entropy is in fact continuous even in the case of a Z^2-action generated by a cellular automation map.

First we embed the lattice Z^2 into the 2-dimensional real vector space R^2. We call a vector $\vec{v}$ a rational vector if the slope of the vector is rational. Otherwise, we call the vector an irrational vector. Let P be a partition of X according to the alphabets. We denote the partition $T^iS^j(P)$ by $P_{i,j}$. Recall the definition of the directional entropy in the direction of $\vec{v}$:

$$h(\vec{v}) = \sup_B \overline{\lim_{t\to\infty}} \frac{1}{t} H\left(\bigvee_{(i,j)\in B+[0,t]\vec{v}} P_{i,j} \right),$$

where $B + [0,t]\vec{v} = \{(i,j) \in Z^2$;there exists $(k,\ell) \in B$ such that $(i,j) - (k,\ell) = \alpha\vec{v}$ for some $\alpha \in [0,t]\}$. Supremum is taken over all bounded subsets of Z^2.

Given a vector $\vec{v} \in R^2$, which is not a scalar multiple of $(1,0)$, we define $w = cot\theta$ where θ is the angle between the vector $\vec{v}$ and x-axis. We note that $(w, 1)$ is a scalar multiple of $\vec{v}$.

It is not difficult to check that

$$h(\vec{v}) = \lim_{n\to\infty} \overline{\lim_{t\to\infty}} \frac{1}{t} H\left(\bigvee_{j=0}^{[ty]} \bigvee_{-n+jw+a<i<n+jw+a} P_{i,j} \right)$$

for any $a \in [0,1]$. Clearly $h(\lambda\vec{v}) = \lambda h(\vec{v})$ and $h(T^pS^q) = h(\vec{v})$ where $\vec{v} = (p,q)$. (see[Mi1])

Also we recall the definition of the cone entropy defined by D.Lind[Li], denoted by $h^c(\vec{v})$ in the direction of $\vec{v} = (x,y)$. Let $\theta_0 = tan^{-1}\frac{y}{x}$. We consider the vector $\vec{v}_\theta = (x_\theta, y)$ and $\vec{v}_{-\theta} = (x_{-\theta}, y)$ where x_θ and $x_{-\theta}$ satisfy $\frac{y}{x_\theta} = \tan(\theta_0 + \theta)$ and $\frac{y}{x_{-\theta}} = \tan(\theta_0 - \theta)$ respectively. Note that $x_{-\theta} \geq x_\theta$. Cone entropy is defined as follows

$$h^c(\vec{v}) = \lim_{\theta \to 0} \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{j=0}^{[ny]} \bigvee_{jx_\theta < i < jx_{-\theta}} P_{i,j}\right)$$

We denote the set $\{(i,j) : 0 \leq j < \infty, -m + jw \leq i \leq m + jw\}$ by $R_m(\vec{v})$ and the set $\{(i,j) : 0 \leq j \leq [ny], -m + jw \leq i \leq m + jw\}$ by $R_m^n(\vec{v})$. If $y = 0$, then $R_m(\vec{v})$ denote the set $\{(i,j) : 0 \leq j \leq m, 0 \leq i < \infty\}$ and $R_m^n(\vec{v})$ denote the set $\{(i,j) : 0 \leq j \leq m, 0 \leq i \leq n\}$. Sometimes by $R_m^n(\vec{v})$ we denote also a parallelogram whose base is not necessarily centered arround (0,0) as long as the length of the base is m and the height is $[ny]$. We will call $C_\theta(\vec{v}) = \{(i,j) : 0 \leq j < \infty, jx_\theta \leq i \leq jx_{-\theta}\}$ the θ-cone arround $\vec{v}$. Note that $C_\theta(\vec{v})$ is an unbounded cone. We will also call $C_\theta^n(\vec{v}) = \{(i,j) : 0 \leq j \leq n[y], jx_\theta \leq i \leq jx_{-\theta}\}$ the n-θ-cone. Sometimes we delete $\vec{v}$ in the notation if the vector is obvious in the context.

We can write

$$h(\vec{v}) = \lim_{m\to\infty} \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{(i,j)\in R_m^{[ny]}(\vec{v})} P_{i,j}\right)$$

and

$$h^c(\vec{v}) = \lim_{\theta\to\infty} \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{(i,j)\in C_\theta^n(\vec{v})} P_{i,j}\right).$$

Note that $h^c(\vec{v}) \geq h(\vec{v})$ for all $\vec{v}$. It is shown that in the case of a Z^2-action generated by a block map (Cellular Automaton), cone entropy in the direction of $\vec{v}$ is the same as the directional entropy $h(\vec{v})$[Pa2]. We have the following Lemma.

LEMMA 1. *$h^c(\vec{v})$ is upper semi-continuous.*

Proof. We let $\{\vec{v}_i = (x_i, y_i)\} \to \vec{v} = (x_0, y_0)$. Given $\epsilon > 0$, there exists $\alpha(\epsilon)$ such that if $0 < \theta \leq \alpha$, then

$$\lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{(i,j)\in C_\theta^n} P_{i,j}\right) < h^c(\vec{v}) + \epsilon.$$

There exists I such that if $i \geq I$, then

$$tan^{-1}\frac{y_0}{x_0} - \alpha < tan^{-1}\frac{y_i}{x_i} < tan^{-1}\frac{y_0}{x_0} + \alpha.$$

Choose α_i such that each α_i-cone around $\vec{v}_i$ for $i \geq I$ is completely contained in the α-cone around $\vec{v}$. Hence

$$\begin{aligned} h^c(\vec{v}_i) &\leq \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{(i,j)\in C_{\alpha_i}^n(\vec{v}_i)} P_{i,j}\right) \\ &\leq \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{(i,j)\in C_\alpha^n(\vec{v}_i)} P_{i,j}\right) \\ &\leq h^c(\vec{v}) + \epsilon. \end{aligned}$$

We have

$$\overline{\lim}\, h^c(\vec{v}_i) \leq h^c(\vec{v})$$

COROLLARY 2. *If $h^c(\vec{v}) = h(\vec{v})$, then $h(\vec{v})$ is upper semi-continuous. In particular, if a Z^2-action is generated by a block map, then the directional entropy is upper semi-continuous.*

Proof. It is known that if a Z^2-action is generated by a block map, then $h^c(\vec{v}) = h(\vec{v})$[Pa2].

LEMMA 3. *Given $\varepsilon > 0$, there exists K such that if $m_1 \geq m_2 \geq K$, then for sufficiently large n,*

$$\frac{1}{n} H\left(\bigvee_{(i,j)\in R_{m_1}^n} P_{i,j}\right) \leq \frac{1}{n} H\left(\bigvee_{(i,j)\in R_{m_2}^n} P_{i,j}\right) + 2\varepsilon.$$

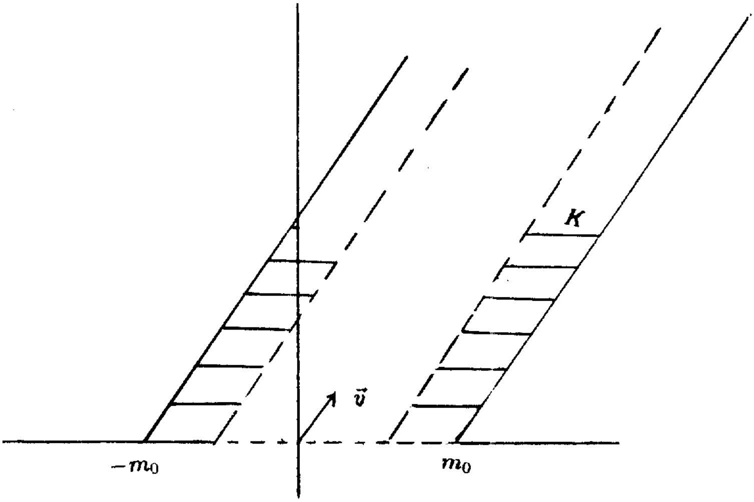
Proof. Clearly we have

$$\frac{1}{n}H\left(\bigvee_{(i,j)\in R^n_{m_1}} P_{i,j}\right) \geq \frac{1}{n}H\left(\bigvee_{(i,j)\in R^n_{m_2}} P_{i,j}\right).$$

If $y \neq 0$, then by scaling, we may assume that $y = 1$. Hence $[ny] = n$. Choose K_0 such that $\sum_{k>K_0} H\left(P_{0,1}|\bigvee_{i=-k}^{k} P_{i,0}\right) < \varepsilon$. Depending on the given vector $\vec{v}$, there exists $K \geq K_0$(independent of m_0) such that for all j,

$$H\left(\bigvee_{-m_0+jx\leq i\leq m_0+jx} P_{i,j}| \bigvee_{-m_0+(j-1)x\leq i\leq m_0+(j-1)x} P_{i,j-1}\right)$$
$$\leq H\left(\bigvee_{-m_0+jx\leq i\leq -m_0+jx+K} P_{i,j} \vee \bigvee_{m_0+jx-K\leq i\leq m_0+jx} P_{i,j}| \bigvee_{-m_0+(j-1)x\leq i\leq m_0+(j-1)x} P_{i,j}\right) + \varepsilon.$$

(See Figure 1)



(Figure 1)

We have

$$\frac{1}{n}H\left(\bigvee_{(i,j)\in R_m^n} P_{i,j}\right)$$
$$=\frac{1}{n}H\left(\bigvee_{-m\le i\le m} P_{i,j}\right)$$
$$+\frac{1}{n}\sum_{j=1}^{n} H\left(\bigvee_{-m+jx\le i\le m+jx} P_{i,j} \mid \bigvee_{(i,\ell)\in R_m^{j-1}} P_{i,\ell}\right).$$

We note that

$$\frac{1}{n}\sum_{j=1}^{n} H\left(\bigvee_{-m+jx\le i\le m+jx} P_{i,j} \mid \bigvee_{(i,\ell)\in R_m^{j-1}} P_{i,\ell}\right)$$
$$=\frac{1}{n}\sum_{j=1}^{n} H\left(\bigvee_{-m+jx\le i\le m+jx+K} P_{i,j} \mid \bigvee_{(i,\ell)\in R_m^{j-1}} P_{i,\ell}\right)$$
$$+\frac{1}{n}\sum_{j=1}^{n} H\left(\bigvee_{-m+jx+K<i\le m+jx-K} P_{i,j} \mid \bigvee_{(i,\ell)\in R_m^{j-1}} P_{i,\ell} \vee \bigvee_{-m+jx\le i\le -m+jx+K} P_{i,j}\right)$$
$$+\frac{1}{n}\sum_{j=1}^{n} H\left(\bigvee_{m+jx-K\le i\le m+jx} P_{i,j} \mid \bigvee_{(i,\ell)\in R_m^{j-1}} P_{i,\ell} \vee \bigvee_{-m+jx\le i\le m+jx-K} P_{i,j}\right).$$

We denote the first, second and the third part of the right hand side of the above equation by $\eta(1,m), \eta(2,m)$ and $\eta(3,m)$ respectively. When we compare $\frac{1}{n}H\left(\bigvee_{(i,j)\in R_{m_1}^n} P_{i,j}\right)$ and $\frac{1}{n}H\left(\bigvee_{(i,j)\in R_{m_2}^n} P_{i,j}\right)$, we have the following.

(i) $\frac{1}{n}H\left(\bigvee_{-m_1\le i\le m_1} P_{i,0}\right) - \frac{1}{n}H\left(\bigvee_{-m_2\le i\le m_2} P_{i,0}\right)$
$\le \frac{2(m_1-m_2)}{n} H(P_{0,0})$.

(ii) By translating the partitions by $T^{-(m_1-m_2)}$, it is easy to see that $\eta(1,m_1) \le \eta(1,m_2)$.

(iii) Similarly it is easy to see that $\eta(3,m_1) \le \eta(3,m_2)$.

(iv) $0 \le \eta(2,m_1), \eta(2,m_2) \le \varepsilon \quad \forall m_1, m_2 \ge K$.

Hence if we choose n large enough so that

$$\frac{2(m_1-m_2)}{n}H(P_{0,0})<\varepsilon,$$

then we have the claim.

If y=0, then clearly we have

$$\begin{aligned}
&\frac{1}{n}H\left(\bigvee_{(i,j)\in R^n_{m_1}} P_{i,j} \mid \bigvee_{(i,j)\in R^n_{m_2}} P_{i,j}\right)\\
\leq&\frac{1}{n}H\left(\bigvee_{\substack{m_2\leq j\leq m_1\\ -m_1\leq j\leq -m_2}}\ \bigvee_{0\leq i<K} P_{i,j}\right)+\varepsilon\\
=&\frac{2K}{n}(m_1-m_2)H(P_{0,0})+\varepsilon<2\varepsilon.
\end{aligned}$$

THEOREM 4. *If S satisfies the condition that*

$$\sum_{k=0}^{\infty}H\left(P_{0,1} \mid \bigvee_{i=-k}^{k} P_{i,0}\right)<\infty,$$

then we have $h^c(\vec{v})=h(\vec{v})$.

Proof. Since the proof is almost the same as that of theorem 1 in [P2], we will briefly sketch the main differences. In the case of $y\neq 0$, by scaling we may assume y=1. Given $\varepsilon>0$, K_0 and K are chosen as in Lemma 3. Note that for all $m\geq K$, we have

$$\begin{aligned}
H^m(\vec{v})&=\lim_{n\to\infty}\frac{1}{n}H\left(\bigvee_{(i,j)\in R^n_m} Pi,j\right)\\
&\leq\lim_{n\to\infty}\frac{1}{n}H\left(\bigvee_{j=0}^{n-1}\ \bigvee_{i=-K+jx}^{K+jx} P_{i,j}\right)+\varepsilon\\
&=H^K(\vec{v})+\varepsilon.
\end{aligned}$$

Since $h(\vec{v})=\lim_{m\to\infty}H^m(\vec{v})$,

$$H^K(\vec{v})\leq h(\vec{v})\leq H^K(\vec{v})+\varepsilon.$$

Since we always have $h^c(\vec{v}) \geq h(\vec{v})$, we need to show that $h^c(\vec{v}) - h(\vec{v})$ is arbitrarily small. Given $\varepsilon > 0$, choose n_0 such that if $n \geq n_0$, then

$$\left|\frac{1}{n} H\left(\bigvee_{(i,j)\in R_K^n} P_{i,j}\right) - h(\vec{v})\right| < 2\varepsilon.$$

We choose θ sufficiently small and $n(\geq n_0)$ sufficiently large so that

(i) $h^c(\vec{v}) - 2\varepsilon \leq \frac{1}{n} H\left(\bigvee_{(i,j)\in C_\theta^n} P_{i,j}\right) \leq h^c(\vec{v}) + 2\varepsilon$

(ii) $(x_{-\theta} - x_\theta) H(P_{0,0}) < \varepsilon$.

(iii) $\frac{1}{n} H\left(\bigvee_{(i,j)\in C_\theta^n} P_{i,j}\right) \geq \frac{1}{n} H\left(\bigvee_{(i,j)\in R_K^n} P_{i,j}\right)$

We compute

$$\begin{aligned}
&h^c(\vec{v}) - h(\vec{v}) \\
&\leq \left(\frac{1}{n} H\left(\bigvee_{(i,j)\in C_\theta^n} P_{i,j}\right) + 2\varepsilon\right) - \left(\frac{1}{n} H\left(\bigvee_{(i,j)\in R_K^n} P_{i,j}\right) - 2\varepsilon\right) \\
&\leq \frac{1}{n}\left(H\left(\bigvee_{(i,j)\in R_L^n} P_{i,j}\right) - H\left(\bigvee_{(i,j)\in R_K^n} P_{i,j}\right)\right) + 4\varepsilon, \\
&\qquad \text{where } L = \tfrac{1}{2}(x_{-\theta} - x_\theta)n \\
&\leq \frac{2}{n}(L - K) H(P_{0,0}) + 4\varepsilon \\
&< \frac{1}{n}(x_{-\theta} - x_\theta) n H(P_{0,0}) + 4\varepsilon \\
&< 5\varepsilon
\end{aligned}$$

Hence we have $h^c(\vec{v}) = h(\vec{v})$.

If y=0, then by choosing θ small enough so that $\frac{1}{tan\theta} > K$, it is easy to see that $h^c(\vec{v}) = h(\vec{v})$.

THEOREM 5. *If S is a finitary code with finite expected code length, then $h(\vec{v})$ is upper semi-continuous.*

Proof. It is enough to show that S satisfies the condition

$$\sum_{k=0}^{\infty} H\left(P_{0,1} \mid \bigvee_{i=-k}^{k} P_{i,0}\right) < \infty.$$

We let $F_k = \{x\colon x_{[-k,k]}$ uniquely determines the 0^{th} coordinate of $Sx\}$ and let $E_k = F_k - \cup_{i=0}^{k-1} F_i$. Note that $E_k' s$ are disjoint and E_k is a union of the atoms of the partition $Q_k = \bigvee_{i=-k}^{k} P_{i,0} = \{Q_{k,0}, \cdots, Q_{k,\ell_k - 1}\}$. Hence F_k is a union of atoms of the partition Q_k. Let L denote the number of atoms of the partition $P_{0,1} = \{R_0, \cdots, R_{L-1}\}$. Since S is a finitary code with finite expected code length, we have

$$\sum_{k=0}^{\infty}(2k+1)\mu E_k < \infty.$$

We compute

$$\begin{aligned}
&\sum_{k=0}^{\infty} H\left(P_{0,1} | \bigvee_{i=-k}^{k} P_{i,0}\right)\\
&= \sum_{k=0}^{\infty} H(P_{0,1}|Q_k)\\
&= \sum_{k=0}^{\infty}\sum_{j=0}^{\ell_k - 1} \mu Q_{k,j}\left(\sum_{i=0}^{L-1}\frac{\mu(R_i \cap Q_{k,j})}{\mu Q_{k,j}}\log\frac{\mu(R_i\cap Q_{k,j})}{\mu Q_{k,j}}\right)\\
&\le \sum_{k=0}^{\infty}\sum_{\substack{j=0\\ Q_{k,j}\not\subset F_k}}^{\ell_k-1} \mu Q_{k,j}(\log L)\\
&= \sum_{k=0}^{\infty}\log L \sum_{\substack{j=0\\ Q_{k,j}\not\subset F_k}}^{\ell_k-1}\mu Q_{k,j}\\
&= \sum_{k=0}^{\infty}\log L(1-\mu F_k)\\
&= \log L\sum_{k=0}^{\infty}(1-\mu F_k)\\
&= \log L\sum_{k=0}^{\infty}(1-\sum_{\ell=0}^{k}\mu E_\ell)\\
&= \log L\left(\lim_{k\to\infty}((k+1)-(k+1)(\mu E_0+\mu E_1+\cdot\cdot+\mu E_k)) + \sum_{\ell=1}^{k-1}\ell\mu E_\ell\right)\\
&= \log L\lim_{k\to\infty}\left(k(1-\mu F_{k-1}) + \sum_{\ell=1}^{k-1}\ell\mu E_\ell\right)
\end{aligned}$$

$$
\begin{aligned}
&= \log L\left(\sum_{\ell=1}^{\infty} \ell\mu E_\ell + \lim_{k\to\infty} k\mu F_{k-1}^c\right) \\
&\le \log L\left(\sum_{\ell=1}^{\infty} \ell\mu E_\ell + \lim_{k\to\infty} k\left(\sum_{\ell=k}^{\infty} \mu E_\ell\right)\right) \\
&= \log L\left(\sum_{\ell=1}^{\infty} \ell\mu E_\ell + \lim_{k\to\infty} \sum_{\ell=k}^{\infty} \ell\mu E_\ell\right) \\
&= \log L\left(\sum_{\ell=1}^{\infty} \ell\mu E_\ell\right) < \infty
\end{aligned}
$$

This completes our proof.

References

[He] G.A.Hedlund, *Endomorphisms and automorphisms of the shift dynamical system,*, Math. Syst.Theor. **3** (1969), 320-375.

[Li] D.Lind, *personal communications.*

[Mi1] J.Milnor, *On the entropy geometry of cellular automata.*, Complex Systems **2** (1988), 357-386.

[Mi2] J.Milnor, *Directional entropies of cellular automation-maps*, Nato ASI Series **F20** (1986), 113-115.

[Pa1] K.K.Park, *Continuity of directional entropy*, to appear in Osaka Journal of Math.

[Pa2] K.K.Park, *Entropy of a skew product with a Z^2-action*, preprint.

[Pa3] K.K.Park, *A counterexample of the etropy of a skew product with a Z^2-action*, preprint.

[Si] Y.Sinai, *An answer to a question by J.Milnor*, Comment. Math. Helv. **60** (1985), 173-178.

[Th] J.P.Thouvenot, *personal communications.*

Department of Mathematics
Ajou University
Suwon 441-749, Korea