Zeitschrift: Commentarii Mathematici Helvetici
Herausgeber: Schweizerische Mathematische Gesellschaft
Band: 60 (1985)

Artikel: An answer to a question by J. Milnor.
Autor: Sinai, Ya. G.
DOI: https://doi.org/10.5169/seals-46307

Comment. Math. Helvetici 60 (1985) 173–178 0010–2571/85/020173–06$01.50 + 0.20/0

An answer to a question by J. Milnor

YA. G. SINAI

We consider two commuting automorphisms T_1, T_2 of the Lebesque space $(M, \mathcal{M}, \mu)$ such that $h_{m,n} = h(T_1^m T_2^n) < \infty$ where h is the measure-theoretic entropy. Under additional assumptions we show the existence of the limits $\lim (1/m) h_{m,n}$ where $m \to \infty$, $n \to \infty$, $m/n \to \omega$ and ω is an irrational number.

§1. Formulation of the problem and the result

Let $X = \{x^{(1)}, \ldots, x^{(\kappa)}\}$ be a finite alphabet and M be the space of double-infinite sequences $x = \{x_n\}_{-\infty}^{\infty}$, $x_n \in X$, S is the shift in M, i.e. $Sx = x' = \{x'_n\}$, $x'_n = x_{n+1}$. Then M is a compact topological space in topology of direct product and S is a homeomorphism of M. Assume that a function $f(x_{-r}, \ldots, x_r)$ with values in X is given. It generates a homomorphism T of M by the formula: $Tx = y = \{y_n\}_{-\infty}^{\infty}$, $y_n = f(x_{n-r}, \ldots, x_{n+r})$. S and T commute and we assume that they generate an action of the group $\mathbb{Z}^2$ on M: for $(m, n) \in \mathbb{Z}^2$ the corresponding transformation is $T_{m,n} = S^m T^n$. The described situation was considered by Professor J. Milnor in his talk "Cellular automata as discrete dynamical systems" during the celebration of the 20-th anniversary of the Forschungsinstitut fur Mathematik, ETH in Zurich. He formulated the following question. Assume that μ is a normed ergodic measure invariant under the action of $\mathbb{Z}^2$. Denote $h_{m,n} = h(S^m T^n)$ measure-theoretic entropy of $T_{m,n}$ with respect to μ. It is easy to show that $h_{m,0} < \infty$ for all $-\infty < m < \infty$. We shall consider the case when $h_{m,n} < \infty$ for all $-\infty < m$, $n < \infty$. From the properties of entropy (see [1]) it follows that the function $h_{m,n}$ is an homogeneous function of the first degree, i.e. $h_{\kappa m, \kappa n} = |\kappa|\, h_{m,n}$. Fix an irrational number $\omega_0 > 0$ and choose a sequence $(m_i, n_i) \in \mathbb{Z}^2$, $m_i \to \infty$, $n_i \to \infty$, $m_i/n_i \to \omega_0$ as $i \to \infty$. The question is whether there exists a limit $\lim_{i\to\infty} (1/\sqrt{m_i^2 + n_i^2}) h_{m_i,n_i}$ which can be called as entropy per unit of length in the direction ω_0. The aim of this paper is to give an affirmative answer to this question. It will be more convenient to show the existence of the limit $\lim_{i\to\infty} (1/n_i) h_{m_i,n_i}$ which is equivalent to the first one.

We introduce the partition ξ into κ sets C_κ, $1 \le i \le \kappa$, $C_i = \{x \mid x_0 = x^{(i)}\}$, $\xi_{m,n} = T_{m,n}\xi$. We shall use later standard notations and facts of the theory of measurable partitions and measure-theoretic entropy (there are many good references, we shall mention only few of them, [1], [2], [3]). By $I = I(a, \omega)$ we denote the segment on the plane joining the points $(a, 0)$ and $(a+\omega^{-1}, 1)$ and $\Gamma(a, \omega)$ is the half-line $y = \omega(x-a)$, $y \leq 1$. It is clear that $I(a, \omega) \subset \Gamma(a, \omega)$. We shall always consider the case $\omega > 0$. The main role in our analysis play the conditional entropies

$$\mathscr{H}_r(I) = H\Big(\bigvee_{m \geq a+\omega^{-1}} \xi_{m,1} \Big| \bigvee_{n=0}^{\infty} \bigvee_{m \geq a+\omega^{-1}n} \xi_{m,-n}\Big)$$

$$\mathscr{H}_l(I) = H\Big(\bigvee_{m \leq a+\omega^{-1}} \xi_{m,1} \Big| \bigvee_{n=0}^{\infty} \bigvee_{m \leq a+\omega^{-1}n} \xi_{m,-n}\Big)$$

$$\mathscr{H}(I) = \mathscr{H}_r(I) + \mathscr{H}_l(I).$$

It is easy to see that both $\mathscr{H}_r(I)$, $\mathscr{H}_l(I)$ are finite. We shall list three properties of them which will be used later:

1. $\mathscr{H}_r(I)$, $\mathscr{H}_l(I)$ are periodic functions of a with the period 1 for each fixed ω;

2. if ω is a rational number, $\omega = p/q$, then $\mathscr{H}_r(I)$, $\mathscr{H}_l(I)$ are constants on each interval of a of the length $1/p$ where the half-lines $\Gamma(a, \omega)$ do not pass through points of the lattice $\mathbb{Z}^2$.

3. if ω is irrational and $\Gamma(a, \omega)$ does not pass through points of the lattice $\mathbb{Z}^2$ then $\mathscr{H}_r(I)$, $\mathscr{H}_l(I)$ are continuous at the point (a, ω).

The last property follows easily from the properties of continuity of conditional entropy. We shall use also a transformation Q in the space of segments $I(a, \omega)$, where $Q(I(a, \omega)) = I(a', \omega)$, $a' = a + \omega^{-1}$.

Our first result is the following theorem.

THEOREM 1. *Let $p > 0$, $q > 0$ have no common factor. Then $h_{p,q} = \sum_{i=0}^{p-1} \mathscr{H}(Q^i(I)) = p \int_0^1 \mathscr{H}(I)\, da$ for any interval $I = I(a, -q/p)$.*

The proof of Theorem 1 is given in §2.

THEOREM 2. *Let ω_0 be an irrational number, (m_i, n_i) be a sequence of points of the lattice $\mathbb{Z}^2$, $m_i, n_i \to \infty$ and $m_i/n_i \to \omega_0$ as $i \to \infty$. Then*

$$\lim_{i\to\infty} \frac{1}{n_i} h_{m_i,n_i} = \int_0^1 \mathscr{H}(I(a, \omega_0))\, da.$$

Proof of Theorem 2. We have from Theorem 1

$$\frac{1}{n_i} h_{m_i,n_i} = \int_0^1 \mathcal{H}(I(a, m_i/n_i))\, da.$$

All functions $\mathcal{H}(I(a, m_i/n_i))$ are uniformly bounded and non-negative. It follows from the property 3 that for almost every a

$$\lim_{i\to\infty} \mathcal{H}(I(a, m_i/n_i)) = \mathcal{H}(I(a, \omega_0)).$$

Thus in view of Lebesgue dominance theorem we have the desired result. Q.E.D.

In §3 we make some additional remarks.

The author expresses his sincere gratitude to Professor Milnor for the formulation of problem and useful discussions. He is also grateful to Professors B. Eckmann, H. Föllmer, J. Fröhlich, K. Hepp, R. Jost, J. Moser, K. Osterwalder and other collegues from ETH, Zurich for their warm hospitality.

§2. Proof of Theorem 1

It follows from the properties of measure-theoretic entropy that

$$h_{q,p} = \lim_{s\to\infty} H\left(\bigvee_{n=1}^{p} \bigvee_{a+\omega^{-1}n-s\le m\le a+\omega^{-1}n+s} \xi_{m,n} \,\middle|\, \bigvee_{n\le 0} \bigvee_{a+\omega^{-1}n-s\le m\le a+\omega^{-1}n+s} \xi_{m,n} \right), \qquad \omega = q/p.$$

The last conditional entropy is equal to

$$\sum_{l=1}^{p} H\left(\bigvee_{a+\omega^{-1}l-s\le m\le a+\omega^{-1}l+s} \xi_{m,l} \,\middle|\, \bigvee_{n<l} \bigvee_{|m-a-\omega^{-1}n|\le s} \xi_{m,n} \right)$$
$$= \sum_{l=1}^{p} H\left(\bigvee_{a+\omega^{-1}l-s\le m\le a+\omega^{-1}l+s} \xi_{m,1} \,\middle|\, \bigvee_{n<0} \bigvee_{|m-a-\omega^{-1}n|\le s} \xi_{m,n} \right).$$

We shall show that the l-th term converges as $s\to\infty$ to $\mathcal{H}(Q^l(I))$. It is sufficient to consider $l=1$, other terms are treated in the same way. From the description of our system it follows easily that

$$H\Big(\bigvee_{a+\omega^{-1}-s\le m\le a+\omega^{-1}+s}\xi_{m,1}\Big|\bigvee_{n\le 0}\ \bigvee_{a+\omega^{-1}n-s\le m\le a+\omega^{-1}n+s}\xi_{m,n}\Big)$$

$$=H\Big(\bigvee_{a+\omega^{-1}-s\le m\le a+\omega^{-1}-s+r}\xi_{m,1}\vee\bigvee_{a+\omega^{-1}+s-r\le m\le a+\omega^{-1}+s}\xi_{m,1}\Big|$$

$$\bigvee_{n\le 0}\ \bigvee_{a+\omega^{-1}n-s\le m\le a+\omega^{-1}n+s}\xi_{m,n}\Big)$$

$$=H\Big(\bigvee_{a+\omega^{-1}+s-r\le m\le a+\omega^{-1}+s}\xi_{m,1}\Big|\bigvee_{n\le 0}\ \bigvee_{a+\omega^{-1}n-s\le m\le a+\omega^{-1}n+s}\xi_{m,n}\Big)$$

$$+H\Big(\bigvee_{a+\omega^{-1}-s\le m\le a+\omega^{-1}-s+r}\xi_{m,1}\Big|\bigvee_{a+\omega^{-1}+s-r\le m\le a+\omega^{-1}+s}\xi_{m,1}$$

$$\times\bigvee_{n\le 0}\ \bigvee_{a+\omega^{-1}n-s\le m\le a+\omega^{-1}n+s}\xi_{m,n}\Big)\quad(1)$$

The first term in (1) is equal to

$$H\Big(\bigvee_{a+\omega^{-1}-[a+\omega^{-1}]-r\le m\le a+\omega^{-1}-[a+\omega^{-1}]}\xi_{m,1}\Big|\bigvee_{n\le 0}\ \bigvee_{a+\omega^{-1}n-2s-[a+\omega^{-1}]\le m\le a+\omega^{-1}n-[a+\omega^{-1}]}\xi_{m,n}\Big)$$

It follows from the properties of continuity of conditional entropy that this expression converges to $\mathcal{H}_l(I)$. We shall show that the second term in (1) converges to $\mathcal{H}_r(I)$. We have

$$H\Big(\bigvee_{a+\omega^{-1}-s\le m\le a+\omega^{-1}-s+r}\xi_{m,1}\Big|\bigvee_{a+\omega^{-1}+s-r\le m\le a+\omega^{-1}+s}\xi_{m,1}$$

$$\vee\bigvee_{n\le 0}\ \bigvee_{a+\omega^{-1}n-s\le m\le a+\omega^{-1}n+s}\xi_{m,n}\Big)$$

$$=H\Big(\bigvee_{a+\omega^{-1}-[a+\omega^{-1}]\le m\le a+\omega^{-1}-[a+\omega^{-1}]+r}\xi_{m,1}\Big|\bigvee_{a+\omega^{-1}+2s-r-[a+\omega^{-1}]\le m\le a+\omega^{-1}-[a+\omega^{-1}]+2s}\xi_{m,1}$$

$$\vee\bigvee_{n\le 0}\ \bigvee_{a+\omega^{-1}(n+1)-[a+\omega^{-1}]\le m\le a+\omega^{-1}(n+1)-[a+\omega^{-1}]+2s}\xi_{m,n}\Big).$$

We denote

$$\eta=\bigvee_{a+\omega^{-1}-[a+\omega^{-1}]\le m\le a+\omega^{-1}-[a+\omega^{-1}]+r}\xi_{m,1}$$

and $C_\eta(x)$ is an element containing $x \in M$. Also let us introduce the partitions

$$\zeta_s = \bigvee_{a+\omega^{-1}-[a+\omega^{-1}]\leq m\leq a+\omega^{-1}-[a+\omega^{-1}]+r+2s} \xi_{m,0}$$

$$\vee \bigvee_{n<0} \bigvee_{a+\omega^{-1}(n+1)-[a+\omega^{-1}]\leq m\leq a+\omega^{-1}(n+1)-[a+\omega^{-1}]+2s} \xi_{m,n},$$

$$\zeta^+ = \bigvee_{n\leq 0} \bigvee_{a+\omega^{-1}(n+1)-[a+\omega^{-1}]\leq m} \xi_{m,n}$$

In view of Doob's theorem on convergence of conditional probabilities

$$\mu(C_\eta(x) \mid C_{\zeta_s}(x)) \to \mu(C_\eta(x) \mid C_{\zeta^+}(x)) \quad \text{a.e.},$$

where $C_{\zeta_s}(x)$, $C_{\zeta^+}(x)$ are elements of corresponding partitions containing x. But

$$\mu\Bigg(C_\eta(x) \Bigg| \bigvee_{a+\omega^{-1}+2s-r-[a+\omega^{-1}]\leq m\leq a+\omega^{-1}-[a+\omega^{-1}]+2s} \xi_{m,1}$$

$$\vee \bigvee_{n\leq 0} \bigvee_{a+\omega^{-1}(n+1)-[a+\omega^{-1}]\leq m\leq a+\omega^{-1}(n+1)-[a+\omega^{-1}]+2s} \xi_{m,n}\Bigg) \qquad (2)$$

can be represented as finite linear combinations of $\mu(C_\eta(x) \mid C_{\zeta_s}(x))$. This shows easily that the conditional probabilities (2) also converge a.e. as $s\to\infty$ to $\mu(C_\eta(x) \mid C_{\zeta^+}(x))$. This gives the desired result. Q.E.D.

§3. Several general remarks

Let us consider two commuting automorphisms T_1, T_2 of Lebesgue space $(M, \mathcal{M}, \mu)$. Then we have a measure-preserving action of the group $\mathbb{Z}^2$ on M and we shall assume that it is ergodic and at least one of automorphisms T_1, T_2 is also ergodic. Without any loss of generality we can assume T_1 is ergodic. If the measure-theoretic entropy $h(T_1)$ is finite one can find a finite generating partition $\xi = \{C_1, \ldots, C_\kappa\}$ in view of Krieger's theorem [4]. It means that T_1 is isomorphic to the shift in the space of doubly-infinite sequences written in the alphabet of κ symbols. If $T_2 x = y = \{y_n\}$ then $y_n = f(x_n, x_{n\pm1}, x_{n\pm2}, \ldots)$ where f is a measurable function with the values in the space $\{1, 2, \ldots, \kappa\}$. Thus the pair (T_1, T_2) is represented as a system of cellular automata but maybe with an infinite memory. Our arguments presented above can be extended to the case when f can be approximated sufficiently well by functions of finite number of variables. However, the general case remains completely open. One can mention also an interesting paper by G. A. Galperin [5] where some results concerning topological entropy of systems of cellular automata were established.

REFERENCES

[1] Rohlin, V. A. *Lectures on the entropy theory of transformations with invariant measure.* Uspekhi Math. Nauk. *22*, No 5, 3–56 (1967).
[2] Parry, W. *Entropy and Generators in Ergodic Theory.* New Haven and London: Yale University Press, 1966.
[3] Cornfeld, I. P., Fomin, S. V. and Sinai, Ya. G. *Ergodic Theory.* Springer-Verlag, Berlin–Heidelberg–New York, 1982.
[4] Krieger, W. *On entropy and generators of measure preserving transformations.* Trans. Amer. Math. Soc., *149*, No 2, 453–464 (1970).
[5] Galperin, G. A. *About an entropy characteristic of homogeneous media with local interaction.* Uspekhi Math. Nauk., *36*, No 6, 207–208 (1981).

L. D. Landau Institute of Theoretical Physics,
Academy of Sciences of USSR,
Moscow, USSR.

Added in proof:

It is clear that theorems 1 and 2 are easily generalized to the case of non-ergodic measures. Indeed, if ν_{inv} is the measurable partition of M into ergodic components of the action of $\mathbb{Z}^2$ corresponding to the measure μ, then

$$\frac{1}{n} h_{m,n} = \int_{M \mid \nu_{\text{inv}}} \frac{1}{n} h(S^m T^n \mid C_{\nu_{\text{inv}}})\, d\mu_{\text{inv}}$$

where μ_{inv} is the induced measure on the factor-space $M \mid \nu_{\text{inv}}$. We showed already for a.e. element $C_{\nu_{\text{inv}}}$ of ν_{inv} the convergence of $(1/n)h(S^m T^n \mid C_{\nu_{\text{inv}}})$, $(m/n) \to \omega$, which implies the convergence of $(1/n)h_{m,n}$.

Also in the same way one can consider the action of the semi-group $\mathbb{Z}^2_+ = \{(m, n): -\infty < m < \infty, n \geqslant 0\}$. In order to get the assertions of theorems 1 and 2 one should replace possible pasts by possible futures in all arguments.