5–10 July 2004, Antalya, Turkey — Dynamical Systems and Applications, Proceedings, pp. 386–394

Generators of Probability Dynamical Systems

M. Ebrahimi
Department of Mathematics, Vali-Asr University, Rafsandjan, Iran
E-mail: mohamad_ebrahimi@hotmail.com

Abstract

One of the important concepts in physics and mathematics is entropy. The concerning results can be improved by dynamical systems and entropy theory techniques. In this paper the concept of entropy will be extended to the countable partitions and we investigate the ergodic properties of probability dynamical systems. In this respect we introduce the generators of probability dynamical systems. A version of Kolmogorov-Sinai theorem concerning the entropy of a probability dynamical system is given.

Introduction and preliminaries

We assume the reader is familiar with the definition of measure [5], dynamical system [9] and ergodic theory [8]. In physics, entropy of a system with a finite number of quantum states is defined as follows:

$$S = -k \sum_{\nu} f_\nu \log(f_\nu); \quad \sum_{\nu} f_\nu = 1,$$

where k is the Boltzmann constant and the sum is over all quantum states.

This formula can be interpreted as a degree of disordering of the system. A system has a unique quantum state with complete ordering and attains zero entropy in this state.

For N different microstates, if $f_\nu = \dfrac{1}{N}$, then

$$S = -k \sum_{\nu=1}^{N} \frac{1}{N} \log \frac{1}{N} = k \log N, .$$

As N becomes larger, the disordering of the system rises up.

The information about the system is proportional to this disordering factor. In mathematics, the entropy of a finite partition, p, of a probability space (X, β, m) is defined as

$$H(p) = -\sum_{i=1}^{n} m(A_i) \log m(A_i),$$

where $p = \{A_1, \ldots, A_n\} \subset \beta$. One can find a nice relation between the mathematical and physical definitions of entropy. In this paper we introduce a definition for the entropy of a countable partition $\xi = \{A_i : i \in \mathbb{N}\}$ and we discuss ergodic theory properties.

1 Entropy of a countable partition of (X, β, m)

1.1. Definition. Let (X, β, m) be a probability space, a partition of (X, β, m) is a disjoint collection of elements of β whose union is X.

We are interested in countable partitions which we denote by Greek letters, *e.g.*, $\xi = \{A_i : i \in \mathbb{N}\}$. We use the convention $\log 0 = 0$.

As in probability theory, we consider a partition $\xi = \{A_i : i \in \mathbb{N}\}$ of (X, β, m) as listing the possible outcomes of an experiment where the probability of the outcome A_i is $m(A_i)$. We associate to this experiment a number $H(\xi)$ which measures the uncertainty removed by performing the experiment mentioned.

1.2. Definition. Let $\xi = \{A_i : i \in \mathbb{N}\}$ be a countable partition of a probability space (X, β, m). The entropy of ξ is defined as

$$H(\xi) = -\log \sup_{i \in \mathbb{N}} m(A_i).$$

1.3. Corollary. *$H(\xi^0) = 0$ where $\xi^0 = \{X, \emptyset\}$, and for each countable partition η with $\eta \neq \xi^0$; $H(\eta) > 0$.*

1.4. Definition. Let $\xi = \{A_i : i \in \mathbb{N}\}$ be a countable partition of a probability space (X, β, m) and let C be a measurable set in β. The conditional entropy of ξ given C is defined by

$$H(\xi|C) = -\log \sup_{i \in \mathbb{N}} m(A_i|C)$$

where

$$m(A_i|C) = \frac{m(A_i \cap C)}{m(C)}, \quad (m(C) \neq 0).$$

1.5. Definition. Two members A, C of β are called independent if $m(A \cap C) = m(A)m(C)$.

If $\xi = \{A_i : i \in \mathbb{N}\}$ is a countable partition of (X, β, m) and C is a measurable set independent of each A_i, we have $H(\xi|C) = H(\xi)$.

1.6. Convention. $H(\xi|\emptyset) = 0$.

1.7. Definition. Let ξ and η be countable partitions of (X, β, m). We say η is a refinement of ξ, denoted by $\xi < \eta$, if each member of ξ is a finite union of some members of η.

1.8. Theorem. *Let $\xi = \{A_i : i \in \mathbb{N}\}$ and $\eta = \{B_j : j \in \mathbb{N}\}$ be two countable partitions of (X, β, m). Then*

i) $\xi < \eta \Longleftrightarrow H(\xi|C) \leq H(\eta|C) \quad \forall C \in \beta$;

ii) $\xi < \eta \Longleftrightarrow H(\xi) \leq H(\eta)$.

Proof. i) Suppose $\xi < \eta$, and then for each $B_j \in \eta$ there exists $A_i \in \xi$ such that $B_j \subseteq A_i$. This implies that $B_j \cap C \subset A_i \cap C$, $\forall C \in \beta$. Then

$$\begin{aligned} m(B_j \cap C) \leq m(A_i \cap C) \quad &\Longrightarrow \quad \sup_{j \in \mathbb{N}} \frac{m(B_j \cap C)}{m(C)} \leq \sup_{i \in \mathbb{N}} \frac{m(A_i \cap C)}{m(C)} \\ &\Longrightarrow \quad \log \sup_{j \in \mathbb{N}} \frac{m(B_j \cap C)}{m(C)} \leq \log \sup_{i \in \mathbb{N}} \frac{m(A_i \cap C)}{m(C)} \\ &\Longrightarrow \quad H(\eta|C) \geq H(\xi|C). \end{aligned}$$

Conversely, if $H(\eta|C) \geq H(\xi|C)$, we have $\sup_{j \in \mathbb{N}} m(B_j \cap C) \leq \sup_{i \in \mathbb{N}} m(A_i \cap C)$ and hence for each $j \in \mathbb{N}$, $B_j \subset A_i$, for some $i \in \mathbb{N}$, *i.e.*, $\xi < \eta$.

ii) It is clear. ■

1.9. Definition. Let $\xi = \{A_i : i \in \mathbb{N}\}$ and $\eta = \{B_j : j \in \mathbb{N}\}$ be countable partitions of (X, β, m). Their joining is the partition

$$\xi \nabla \eta = \{A_i \cap B_j : i, j \in \mathbb{N}\}$$

with lexicographic ordering. If $C \in \beta$, then $\xi \nabla C = \{A_i \cap C : i \in \mathbb{N}\}$.

1.10. Corollary.

i) $H(\xi \nabla \eta) \geq H(\xi)$, $H(\xi \nabla \eta) \geq H(\eta)$ *and if* ξ *and* η *are independent, then* $H(\xi \nabla \eta) = H(\xi) + H(\eta)$;

ii) $H(\xi \nabla C) \geq H(\xi), \quad \forall C \in \beta$;

iii) $H(\xi \nabla C) \geq H(\xi|C) \quad \forall C \in \beta$;

iv) $\xi < \eta \Longleftrightarrow H(\eta \nabla C) \geq H(\xi \nabla C)$, $\forall C \in \beta$.

Proof. i) If ξ and η are independent, then

$$\begin{aligned} H(\xi \nabla \eta) &= -\log \sup_{i,j \in \mathbb{N}} m(A_i \cap B_j) \\ &= -\log \sup_{i,j \in \mathbb{N}} m(A_i) m(B_j) \\ &= -\log \sup_{i \in \mathbb{N}} m(A_i) \sup_{j \in \mathbb{N}} m(B_j) \\ &= -\log \sup_{i \in \mathbb{N}} m(A_i) + (-\log \sup_{j \in \mathbb{N}} m(B_j)) \\ &= H(\xi) + H(\eta). \end{aligned}$$

■

1.11. Definition. Let (X, β, m) be as above. We define the entropy function of a measurable set C by

$$H(C) = -\log m(C).$$

1.12. Corollary. *Let ξ be a countable partition of (X, β, m) and $C, D \in \beta$. Then*
i) $H(C) \geq 0$;
ii) $H(\xi \nabla C) \geq H(C)$;
iii) $D \subseteq C \Longleftrightarrow H(D) \geq H(C)$;
iv) $D \subseteq C \Longleftrightarrow H(\xi \nabla D) \geq H(\xi \nabla C)$;
iv) $H(\xi \nabla C) = H(\xi | C) + H(C)$.

1.13. Definition. If ξ is a countable partition of (X, β, m), the diameter of ξ is defined as follows:

$$\operatorname{diam} \xi = \sup_{A_i \in \xi} m(A_i).$$

1.14. Definition. Let $\xi = \{A_i : i \in \mathbb{N}\}$ and $\eta = \{B_j : j \in \mathbb{N}\}$ be two countable partitions of (X, β, m). The conditional entropy of ξ given η is defined as:

$$H(\xi|\eta) = -\log \sup_{i \in \mathbb{N}} \frac{\operatorname{diam}(A_i \nabla \eta)}{\operatorname{diam} \eta} = -\log \sup_{j \in \mathbb{N}} \frac{\operatorname{diam}(\xi \nabla B_j)}{\operatorname{diam} \eta}.$$

Since $\xi^0 = \{X, \emptyset\}$ represents the outcome of the trivial experiment, $H(\xi|\xi^0) = H(\xi)$ and $H(\xi^0|\xi) = 0$ where $\xi \neq \xi^0$.

1.15. Proposition. *If ξ, η and ζ are countable partitions of a probability space (X, β, m), then*
i) $H(\xi|\eta) \geq 0$;
ii) If $\eta < \zeta$, *then* $H(\xi|\eta) \leq H(\xi \nabla \zeta)$, *especially* $H(\xi|\eta) \leq H(\xi \nabla \eta)$;
iii) If $\xi < \eta$, *then* $H(\xi|\zeta) \leq H(\eta|\zeta)$;
iv) *If ξ and $\eta \nabla \zeta$ are independent, then*

$$H(\xi \nabla \eta | \zeta) = H(\xi) + H(\eta|\zeta).$$

Proof. iii) Since $\xi < \eta$, for each $B_j \in \eta$, there exists $A_i \in \xi$ such that $B_j \subseteq A_i$. Then $B_j \cap C_k \subseteq A_i \cap C_k$, $\forall C_k \in \zeta$. Therefore

$$m(B_j \cap C_k) \leq m(A_i \cap C_k) \quad \forall C_k \in \zeta$$

$$\Longleftrightarrow \quad \frac{\sup_{j \in \mathbb{N}} m(B_j \cap C_k)}{\sup_{k \in \mathbb{N}} m(C_k)} \leq \frac{\sup_{i \in \mathbb{N}} m(A_i \cap C_k)}{\sup_{k \in \mathbb{N}} m(C_k)} \quad \forall C_k \in \zeta$$

$$\Longleftrightarrow \quad H(\xi|\zeta) \leq H(\eta|\zeta).$$

Now we prove (iv):

$$\begin{aligned}
H(\xi\nabla\eta|\zeta) &= -\log\sup_{k\in\mathbb{N}}\frac{\operatorname{diam}(\xi\nabla\eta\nabla C_k)}{\operatorname{diam}\zeta}\\
&= -\log\sup_{k\in\mathbb{N}}\frac{\sup\limits_{i,j\in\mathbb{N}} m(A_i\cap B_j\cap C_k)}{\operatorname{diam}\zeta}\\
&= -\log\sup_{k\in\mathbb{N}}\frac{\sup\limits_{i\in\mathbb{N}} m(A_i)\sup\limits_{j\in\mathbb{N}}(B_j\cap C_k)}{\operatorname{diam}\zeta}\\
&= -\log\sup_{i\in\mathbb{N}} m(A_i)\sup_{k\in\mathbb{N}}\frac{\sup\limits_{j\in\mathbb{N}} m(B_j\cap C_k)}{\operatorname{diam}\zeta}\\
&= -\log\sup_{i\in\mathbb{N}} m(A_i) - \log\sup_{k\in\mathbb{N}}\frac{\sup\limits_{j\in\mathbb{N}} m(B_j\cap C_k)}{\operatorname{diam}\zeta}\\
&= H(\xi)+H(\eta|\zeta).
\end{aligned}$$

■

2 Entropy of a measure-preserving transformation

2.1. Definition. Suppose (X_1,β_1,m_1) and (X_2,β_2,m_2) are probability spaces. A transformation $T: X_1 \longrightarrow X_2$ is measure-preserving if
i) $T^{-1}\beta_2 \subset \beta_1$;
ii) $m_1(T^{-1}B_2) = m_2(B_2), \quad \forall B_2 \in \beta_2$.
2.2. Definition. Suppose $T: X \longrightarrow X$ is a measure-preserving transformation of the probability space (X,β,m). If ξ is a countable partition of (X,β,m), we define the entropy of T with respect to ξ as

$$h(T,\xi) = \lim_{n\to+\infty}\frac{1}{n}H(\nabla_{i=0}^{n-1}T^{-i}\xi)$$

where $T^{-i}\xi = \{T^{-i}A_j : j \in \mathbb{N}\}$, for $\xi = \{A_j : j \in \mathbb{N}\}$, a countable partition of (X,β,m).

To see that $h(T,\xi)$ always exists, let $a_n = H(\nabla_{i=0}^{n-1}T^{-i}\xi) \geq 0$. Then

$$a_{n+p} = H(\nabla_{i=0}^{n+p-1}T^{-i}\xi) \leq H(\nabla_{i=0}^{n-1}T^{-i}\xi) + H(\nabla_{i=n}^{n+p-1}T^{-i}\xi) = a_n + a_p.$$

So $a_{n+p} \leq a_n + a_p$, $\forall n,p$. Hence $\lim\limits_{n\to\infty}\dfrac{a_n}{n}$ exists.
2.3. Theorem. *Let ξ, η and ζ be countable partitions of the probability space (X,β,m). Let T be a measure preserving transformation of (X,β,m). Then the*

following properties hold:
i) $h(T,\xi) \le h(\xi)$;
ii) $\xi < \eta \Longrightarrow h(T,\xi) \le h(T,\eta)$;
iii) $h(T, T^{-1}\xi) = h(T,\xi)$;
iv) $h(T, \nabla_{i=0}^{r-1} T^{-i}\xi) = h(T,\xi)$, $\forall r \ge 1$.

Proof. See [1]. ■

2.4. Theorem. *Let T be a measure-preserving transformation of the probability space (X,β,m). Then*
i) For $k>0$, $h(T^k) = kh(T)$;
ii) If T is invertible, then $h(T^k) = |k|h(T)$, $\forall k \in \mathbb{Z}$.

Proof. i) $\lim\limits_{n\to\infty} \dfrac{1}{n} H(\nabla_{j=0}^{n-1} T^{-kj}(\nabla_{i=0}^{k-1} T^{-i}\xi)) = \lim\limits_{n\to\infty} \dfrac{k}{nk} H(\nabla_{i=0}^{nk-1} T^{-i}\xi) = kh(T,\xi)$.
Then

$$h(T^k, \nabla_{i=0}^{k-1} T^{-i}\xi) = kh(T,\xi).$$

We have

$$\lim_{n\to\infty} \frac{1}{n} H(\nabla_{j=0}^{n-1} T^{-k}(\nabla_{i=0}^{k-1} T^{-i}\xi)) = \lim_{n\to+\infty} \frac{k}{nk} H(\nabla_{i=0}^{nk-1} T^{-i}\xi) = kh(T,\xi).$$

Thus

$$\begin{aligned} kh(T) &= k \sup_{\xi \text{ countable}} h(T,\xi) = \sup_{\xi} h(T^{-k}, \nabla_{i=0}^{k-1} T^{-i}\xi) \\ &= \sup_{\eta \text{ countable}} h(T^k, \eta) = h(T^k). \end{aligned}$$

Also $h(T^k,\xi) \le h(T^k, \nabla_{i=0}^{k-1} T^{-i}\xi) = kh(T,\xi)$, and so $h(T^k) \le kh(T)$.

ii) We show that $h(T^{-1},\xi) = h(T,\xi)$ for each countable partition ξ. But

$$\begin{aligned} h(T^{-1},\xi) &= \lim_{n\to\infty} \frac{1}{n} H(\nabla_{i=0}^{n-1} T^{i}\xi) = \lim_{n\to\infty} \frac{1}{n} H(T^{-(n-1)}(\nabla_{i=0}^{n-1} T^{i}\xi)) \\ &= \lim_{n\to\infty} \frac{1}{n} H(\nabla_{j=0}^{n-1} T^{-j}\xi) \\ &= h(T,\xi). \end{aligned}$$ ■

3 Entropy and generators of probability dynamical systems

3.1. Definition. A probability dynamical system is a complex (X,β,m,T) where (X,β,m) is a probability space and $T : X \longrightarrow X$ is a measure-preserving transformation.

Remark. The probability dynamical system is a discrete-time system with $T^i : X \longrightarrow X$, $i = 1, 2, 3, \ldots$, considered as $T \circ T \circ \cdots \circ T$, i-times. Consequently $T^0 = id$ and $T^{i+j} = T^i \circ T^j$, $\forall i, j \in \mathbb{N}$.

3.2. Definition. Let (X, β, m, T) be a probability dynamical system. The entropy of the probability dynamical system (X, β, m, T) is defined as

$$h(T) = \sup_{\xi} h(T, \xi)$$

where the supremum is taken over all countable partitions of (X, β, m, T).

Let (X_i, β_i, m_i, T_i), $i = 1, 2$ be a probability dynamical system.

The probability dynamical system (X_2, β_2, m_2, T_2) is said to be a homomorphic image of the probability dynamical system (X_1, β_1, m_1, T_1) if there exists a measure-preserving transformation $f : X_1 \longrightarrow X_2$ such that $f \circ T_1 = T_2 \circ f$.

(X_1, β_1, m_1, T_1) and (X_2, β_2, m_2, T_2) are called isomorphic if
i) (X_2, β_2, m_2, T_2) is a homomorphic image of (X_1, β_1, m_1, T_1) under a transformation f.
ii) f is invertible and (X_1, β_1, m_1, T_1) is a homomorphic image of (X_2, β_2, m_2, T_2) under the transformation f^{-1}.

3.3. Theorem. *If (X_i, β_i, m_i, T_i), $i = 1, 2$, are isomorphic dynamical systems, then $h(T_1) = h(T_2)$.*

Proof. Since (X_i, β_i, m_i, T_i) are isomorphic, there exists an invertible measure preserving transformation $f : X_1 \longrightarrow X_2$ such that $f \circ T_1 = T_2 \circ f$.

Let η be any countable partition of (X_2, β_2, m_2, T_2), we get

$$\begin{aligned}
h(T, \eta) &= \lim_{n\to\infty} \frac{1}{n} H(\nabla_{i=0}^{n-1} T_2^{-i} \eta) \\
&= \lim_{n\to\infty} \frac{1}{n} H(f^{-1}(\nabla_{i=0}^{n-1} T_2^{-i} \eta)) \\
&= \lim_{n\to\infty} \frac{1}{n} H(\nabla_{i=0}^{n-1} f^{-1}(T_2^{-i} \eta)) \\
&= \lim_{n\to\infty} \frac{1}{n} H(\nabla_{i=0}^{n-1} T_1^{-i}(f^{-1} \eta)) \\
&= h(T_1, f^{-1}\eta).
\end{aligned}$$

Then $h(T_2) = \sup_{\eta} h(T_2, \eta) = \sup_{\eta} h(T_1, f^{-1}\eta) \leq \sup_{\xi} h(T_1, \xi) = h(T_1)$ where ξ is any countable partition of (X_1, β_1, m_1, T_1).

Similarly $h(T_1) \leq h(T_2)$ and hence $h(T_1) = h(T_2)$. ∎

3.4. Definition. Let (X, β, m) be a probability space and $T : X \longrightarrow X$ be an invertible measure-preserving transformation. A countable partition ξ of (X, β, m)

is said to be a generator of the probability dynamical system (X, β, m, T) if there exists an integer $r > 0$ such that

$$\eta < \nabla_{i=0}^{r} T^{-i}\xi$$

for each countable partition η of (X, β, m).

3.5. Theorem. *If ξ is a generator of the probability dynamical system (X, β, m, T), then*

$$h(T, \eta) \leq h(T, \xi)$$

for each countable partition η of (X, β, m).

Proof. Let η be any countable partition of (X, β, m). Since ξ is a generator, $\eta < \nabla_{i=0}^{r} T^{-i}\xi$.

From Theorem 2.3, $h(T, \eta) \leq h(T, \nabla_{i=0}^{r} T^{-i}\xi) = h(T, \xi)$. ■

Now we can deduce the following version of Kolmogorov-Sinai theorem.

3.6. Theorem. *If ξ is a generator of the probability dynamical system (X, β, m, T), then*

$$h(T) = h(T, \xi).$$

Proof. Obvious. ■

3.7. Theorem. *Consider the transformation $T : X \longrightarrow X$; $T = id$. Then $h(T) = 0$.*

Proof. Since $T = id$, $T^{-1}\xi = \{T^{-1}A_i : i \in \mathbb{N}\} = \{A_i : i \in \mathbb{N}\} = \xi$ for each countable partition ξ of (X, β, m). Hence $T^{-i}\xi = \xi$, $i = 0, 1, 2, 3, \ldots$, and

$$\begin{aligned} h(T, \eta) &= \lim_{n\to\infty} \frac{1}{n} H(\nabla_{i=0}^{n-1} T^{-i}\xi) \\ &= \lim_{n\to\infty} \frac{1}{n} H(\nabla_{i=0}^{n-1} \xi) \\ &= \lim_{n\to\infty} \frac{1}{n} H(\xi) \\ &= 0 \end{aligned}$$

for each countable partition ξ of (X, β, m). It follows that

$$h(T) = \sup_{\xi \text{ countable}} h(T, \xi) = 0. \qquad ■$$

3.9. Corollary. *If $T : X \longrightarrow X$ is a measure-preserving transformation on the probability space (X, β, m) with $T^k = id$ for some $k \neq 0$, then $h(T) = 0$.*

Proof. Since $T^k = id$, $h(T^k) = 0$, and $h(T) = \dfrac{1}{|k|} h(T^k) = 0$. ■

References

[1] Cornfeld I. P., Fomin S. V. and Sinai Ya. G., *Ergodic Theory*, Springer–Verlag, New York, Heidelberg, Berlin, 1980.

[2] Kolmogorov A. N., *A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces*, Dokl. Akad. Nauk SSSR, **119** (1958), 861–864.

[3] Ornstein D. S., *Ergodic Theory, Randomness and Dynamical Systems*, Yale Univ. Press, New York, Haven, 1973.

[4] Parthasarathy K. R., *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.

[5] Parthasarathy K. R., *Introduction To Probability and Measures*, McMillan, London, 1977.

[6] Pollicott M. and Yuri M., *Dynamical Systems and Ergodic Theory*, Cambridge Univ. Press, 1979.

[7] Sinai J. G., *On the notion of entropy of a dynamical system*, Dokl. Akad. Nauk. SSSR, **124** (1959), 768–771.

[8] Sinai Ya. G., *Introduction To Ergodic Theory*, Princeton Univ. Press, Princeton, N. J., 1976.

[9] Sinai Ya. G., *Dynamical Systems*, Vol. 1, Aarhus Univ., 1970.