A Bilateral Version of the Shannon-McMillan-Breiman Theorem

Pierre Tisseur
Laboratoire Génome et Informatique
Université d'Evry, Tour Evry 2.

December 5, 2003

Abstract

We give a new version of the Shannon-McMillan-Breiman theorem in the case of a bijective action. For a finite partition α of a compact set X and a measurable action T on X, we denote by $C^T_{n,m,\alpha}(x)$ the element of the partition $\alpha \vee T^1\alpha \vee \ldots \vee T^m\alpha \vee T^{-1}\alpha \vee \ldots \vee T^{-n}\alpha$ which contains a point x. We prove that for μ-almost all x,

$$\lim_{n+m\to\infty} \left(\frac{-1}{n+m}\right) \log \mu(C^T_{n,m,\alpha}(x)) = h_\mu(T,\alpha),$$

where μ is a T-ergodic probability measure and $h_\mu(T,\alpha)$ is the metric entropy of T with respect to the partition α.

1 Introduction

The Shannon-McMillan-Breiman theorem [2], [3] is used in many problems related to the metric entropy map of an ergodic measure. We extend this well-known result to the case of a bijective dynamical system. Our proof follows the line of Petersen's proof [3]. We illustrate this new result with an example that gives an inequality between shifts and cellular automata entropies and some analog of the Lyapunov exponents. Our bilateral version of the Shannon-McMillan-Breiman theorem is expected to be useful in other areas of dynamical systems.

2 Background material

Let X be a compact space, μ a probability measure on X and T a measurable map from X to X. We denote by α a finite partition of X and by $C^T_{n,\alpha}(x)$ the element of the partition $\alpha \vee T^{-1}\alpha \vee \ldots \vee T^{-n}\alpha$ which contains the point x. For all point x the information map I is defined by

$$I(\alpha)(x) = -\log \mu(C^T_\alpha(x)) = \sum_{A \in \alpha} -\log \mu(A)\chi_A(x),$$

where $C^T_\alpha(x)$ is the element of α which contains x and χ_A is the characteristic function defined by

$$\chi_A(x) = \begin{cases} 1 \text{ if } x \in A, \\ 0 \text{ otherwise.} \end{cases}$$

The information map satisfies

$$\begin{aligned} I(\alpha \vee \beta) &= I(\alpha) + I(\beta|\alpha), & (1) \\ I(T\alpha|T\beta) &= I(\alpha|\beta) \circ T^{-1}. & (2) \end{aligned}$$

These two properties are easily proved from the definition of I and the fact that T is a surjective map. We refer to [3, p. 238], [4, Chap.8] for a detailed proof of (1Background materialequation.1) and (2Background materialequation.1).

A simple formulation of the metric entropy with respect to the partition α is given by

$$h_\mu(T, \alpha) = \lim_{n \to \infty} \int_X I(\alpha| \vee_{k=1}^n T^k\alpha)(x)\, d\mu(x),$$

where

$$I(\alpha|\beta)(x) = -\sum_{A \in \alpha, B \in \beta} \chi_{A \cap B}(x) \log\left(\frac{\mu(A \cap B)}{\mu(B)}\right)$$

is the conditional information map representing the quantity of information given by the partition α knowing the partition β about the point x.

We recall the Shannon-McMillan-Breiman theorem [2] [3].

Theorem 1 (Shannon-McMillan-Breiman's theorem) *If μ is a T-ergodic measure, then for μ-almost all x in a compact X we have*

$$\lim_{n \to \infty} \frac{-1}{n} \log \mu(C^T_{n,\alpha}(x)) = h_\mu(T, \alpha).$$

3 A bilateral version of Shannon-McMillan-Breiman's theorem

In order to prove the main result (Theorem 2) we need to expose two technical lemmas. The proof of Lemma 2 and Theorem 2 requires a bilateral version of the Birkhoff pointwise ergodic theorem: for a T-ergodic measure μ one has

$$\lim_{n+m\to\infty} \frac{1}{n+m+1} \sum_{k=-m}^{n} f\circ T^k(x) = \int_X f(x)d\mu(x)$$

for almost all x with a map f in L_1. This result is easily deduced from the Birkhoff pointwise ergodic theorem (see [4, Chap.10]) by breaking up the infinite sum in two proportional parts.

Lemma 1 *For all integers m and n we have*

$$\sum_{k=-m}^{n-1} I(\alpha|\vee_{j=1}^{n-k} T^{-j}\alpha)\circ T^k = I(\vee_{j=-n}^{m} T^j\alpha) - I(T^{-n}\alpha).$$

Proof. Note that $I(\vee_{j=-n}^{m} T^j\alpha) = I(\vee_{j=0}^{m+n} T^{m-j}\alpha)$. Using (1Background materialequation.1) we have

$$I(\vee_{j=0}^{m+n} T^{m-j}\alpha) = I(\vee_{j=1}^{m+n} T^{m-j}\alpha) + I(T^m\alpha|\vee_{j=1}^{m+n} T^{m-j}\alpha)$$

and from (2Background materialequation.1) we get

$$I(T^m\alpha|\vee_{j=1}^{m+n} T^{m-j}\alpha) = I(\alpha|\vee_{j=1}^{m+n} T^{-j}\alpha)\circ T^{-m}.$$

Hence,

$$I(\vee_{j=-n}^{m} T^j\alpha) = I(\vee_{j=1}^{m+n} T^{m-j}\alpha) + I(\alpha|\vee_{j=1}^{m+n} T^{-j}\alpha)\circ T^{-m}.$$

The same operations on $I(\vee_{j=1}^{m+n} T^{m-j}\alpha)$ yields

$$\begin{aligned} I(\vee_{j=1}^{m+n} T^{m-j}\alpha) &= I(\vee_{j=2}^{m+n} T^{m-j}\alpha) + I(\alpha|\vee_{j=2}^{m+n} T^{1-j}\alpha)\circ T^{-m+1} \\ &= I(\vee_{j=1}^{m+n-1} T^{m-1-j}\alpha) + I(\alpha\vee_{j=1}^{m+n-1} T^{-j}\alpha)\circ T^{-m+1}. \end{aligned}$$

Hence,

$$\begin{aligned} I(\vee_{j=-n}^{m} T^j\alpha) &= I(\vee_{j=1}^{m+n-1} T^{m-1-j}\alpha) \\ &\quad + I(\alpha\vee_{j=1}^{m+n-1} T^{-j}\alpha)\circ T^{-m+1} + I(\alpha|\vee_{j=1}^{m+n} T^{-j}\alpha)\circ T^{-m}. \end{aligned}$$

Iterating similarly $t-1$ times on $I(\vee_{j=1}^{m+n-1} T^{m-1-j}\alpha)$ leads to

$$I(\vee_{j=-n}^{m} T^j\alpha) = I(\vee_{j=1}^{m+n-t} T^{m-t-j}\alpha) + \sum_{k=0}^{t} I(\alpha| \vee_{j=1}^{m+n-k} T^{-j}\alpha)\circ T^{-m+k}.$$

Taking $t = m+n-1$ gives

$$I(\vee_{j=-n}^{m} T^j\alpha) = +I(T^{-n}\alpha) + \sum_{k=0}^{m+n-1} I(\alpha| \vee_{j=1}^{m+n-k} T^{-j}\alpha)\circ T^{-m+k}$$

which completes the proof.

□

Lemma 2 *If μ is a T ergodic measure then for almost all x in X,*

$$\lim_{m+n\to\infty} \frac{1}{m+n+1} \sum_{k=-m}^{n-1} \lim_{s\to\infty} I(\alpha| \vee_{j=1}^{s} T^{-j}\alpha)\circ T^k(x) \quad = \quad \lim_{m+n\to\infty} \frac{1}{m+n+1} \sum_{k=-m}^{n-1} I(\alpha| \vee_{j=1}^{n-k} T^{-j}\alpha)\circ T^k(x).$$

Proof. For notational convenience, we introduce

$$f = \lim_{s\to\infty} I(\alpha| \vee_{j=1}^{s} T^{-j}\alpha) \quad \text{and} \quad F_N = \sup_{s\geq N} |I(\alpha| \vee_{j=1}^{s} T^{-j}\alpha) - f|.$$

It is well known that the sequence $(I(\alpha| \vee_{j=1}^{s} T^{-j}\alpha))_{s\in\mathbb{N}}$ converge almost everywhere and in L_1. The proof of this convergence (see [4, Chap.8] and [3, p.262]) requires the increasing martingale theorem.

We need to show that

$$\lim_{m+n\to\infty} \frac{1}{m+n+1} \sum_{k=-m}^{n-1} \left|I(\alpha| \vee_{j=1}^{n-k} T^{-j}\alpha)\circ T^k - f\circ T^k\right| = 0.$$

Note that

$$\frac{1}{m+n+1} \sum_{k=-m}^{n-1} |I(\alpha| \vee_{j=1}^{n-k} T^{-j}\alpha)\circ T^k - f\circ T^k| \quad \leq \quad \frac{1}{m+n+1} \sum_{k=n-N}^{n-1} |I(\alpha| \vee_{j=1}^{n-k} T^{-j}\alpha)\circ T^k - f\circ T^k| + \frac{1}{m+n+1} \sum_{k=-m}^{n-N-1} |F_N|\circ T^k.$$

If we fix N and let $n+m$ tend to infinity then the first term in the right-hand side of the above inequality goes to zero. Since the map F_N belongs to L_1 (see [3]), the bilateral version of Birkhoff's ergodic theorem applies and we can assert that

$$\lim_{m+n\to\infty} \frac{1}{m+n+1} \sum_{k=-m}^{n-N-1} |F_N| \circ T^k = \int_X F_N \, d\mu.$$

Since $\lim_{N\to\infty} F_N = 0$, the dominated convergence theorem implies that $\int_X F_N \, d\mu$ tends to zero which completes the proof.

□

Theorem 2 *For a bijective map T from X to X and a T-ergodic measure μ, we have for μ-almost all x*

$$\lim_{n+m\to\infty} \frac{-1}{n+m} \log \mu(C^T_{n,m,\alpha}(x)) = h_\mu(T, \alpha),$$

where $C^T_{n,m,\alpha}(x)$ represents the element of the partition $\alpha \vee T\alpha \ldots \vee T^m\alpha \vee T^{-1}\alpha \vee \ldots \vee T^{-n}\alpha$ containing the point x.

Proof. Since the sequence $\left(I(\alpha| \vee_{j=1}^s T^{-j}\alpha)\right)_{s\in\mathbb{N}}$ converges to a L_1 map and by using the dominated convergence theorem, it follows that

$$h_\mu(T,\alpha) = \lim_{s\to\infty} \int_X I(\alpha| \vee_{j=1}^s T^{-j}\alpha)(x) \, d_\mu(x) = \int_X \lim_{s\to\infty} I(\alpha| \vee_{j=1}^s T^{-j}\alpha)(x) \, d_\mu(x)$$

for μ-almost all x. The bilateral version of Birkhoff's ergodic theorem implies that for almost all x

$$h_\mu(T,\alpha) = \lim_{m+n\to\infty} \frac{1}{m+n+1} \sum_{k=-m}^{n-1} \lim_{s\to\infty} I(\alpha| \vee_{j=1}^s T^{-j}\alpha) \circ T^k(x).$$

From Lemma 2 it follows that

$$h_\mu(T,\alpha) = \lim_{m+n\to\infty} \frac{1}{m+n+1} \sum_{k=-m}^{n-1} I(\alpha| \vee_{j=1}^{n-k} T^{-j}\alpha) \circ T^k(x).$$

Using Lemma 1 for almost all x we obtain

$$h_\mu(T,\alpha) = \lim_{m+n\to\infty} \frac{1}{m+n+1} \left(I(\vee_{j=-n}^m T^j\alpha)(x) - I(T^{-n}\alpha)(x)\right).$$

Since $(I(T^{-n}\alpha))_{n\in\mathbf{N}}$ is bounded for μ-almost all point x, the sequence $\frac{I(T^{-n}\alpha)}{n+m+1}$ tends almost surely to zero. Hence,

$$\begin{aligned} h_\mu(T,\alpha) &= \lim_{m+n\to\infty} \frac{1}{m+n+1} I(\vee_{j=-n}^{m} T^j\alpha)(x) \\ &= -\lim_{m+n\to\infty} \frac{1}{m+n} \log \mu(C^T_{n,m,\alpha}(x)). \qquad \Box \end{aligned}$$

4 An illustration

In this example we do not give a definition of the particular discrets dynamical systems called cellular automata; the reader can find a survey in [1] and the complete proof of this illustration in [5]. The bilateral version of the Shannon-McMillan-Breiman theorem is needed to establish an inequality between the entropy of a cellular automaton F denoted by $h_\mu(F,\alpha)$, the entropy $h_\mu(\sigma)$ of a particular bijective cellular automaton σ called the shift, and some discret analog of the Lyapunov exponents. Here the measure μ is F-invariant and σ-ergodic.

The standart Shannon-McMillan-Breiman theorem [4, Chap.10] says that in the case of an invariant measure μ.

$$h_\mu(F,\alpha) = \int \lim_{n\to\infty} \frac{-1}{n} \log \mu\left(C^F_{n,\alpha}(x)\right).$$

In [5] one proves that there exists some integer and bounded maps f_n and g_n such that, for all point x, one has

$$C^F_{n,\alpha}(x) \supset C^\sigma_{f_n(x),g_n(x),\alpha}(x)$$

with $\lim_{n\to\infty} f_n(x)+g_n(x) = +\infty$ for μ-almost all point x for a certain class of cellular automata. For those that do not belong to this class, the entropy is equal to zero (see [5]). With these properties we obtain

$$h_\mu(F,\alpha) \le \int \lim_{n\to\infty} \frac{-1}{n} \log \mu\left(C^\sigma_{f_n(x),g_n(x),\alpha}(x)\right) d\mu(x)$$

and

$$\begin{aligned} h_\mu(F,\alpha) \le \int \liminf_{n\to\infty} & \frac{-1}{f_n(x)+g_n(x)+1} \log \mu\left(C^\sigma_{f_n(x),g_n(x),\alpha}(x)\right) \\ & \times \frac{g_n(x)+f_n(x)+1}{n} d\mu(x). \end{aligned}$$

The bilateral version of the Shannon-McMillan-Breiman theorem implies that

$$h_\mu(F,\alpha) \le h_\mu(\sigma,\alpha) \int \liminf_{n\to\infty} \frac{f_n(x)+g_n(x)+1}{n}.$$

Using the Fatou lemma, we have

$$h_\mu(F,\alpha) \le h_\mu(\sigma,\alpha) \times (\lambda_\mu^+ + \lambda_\mu^-),$$

where

$$\lambda_\mu^+ = \liminf \int \frac{f_n(x)}{n} d\mu(x) \quad \text{and} \quad \lambda_\mu^- = \liminf \int \frac{g_n(x)}{n} d\mu(x)$$

are called the left and right average Lyapunov exponents.

References

[1] François Blanchard, Petr Kurka, and Alejandro Maass, *Topological and measure-theoretic properties of one-dimensional cellular automata*, Phys. D, 103 (1997), pp. 86–99. Lattice dynamics (Paris, 1995).

[2] Leo Breiman, *The individual ergodic theorem of information theory*, Ann. Math. Stat., 28 (1957), pp. 809–811.

[3] Karl Petersen, *Ergodic theory*, Cambridge Studies in Advanced Mathematics, 2, Cambridge University Press, Cambridge, UK, 1983.

[4] Mark Pollicott and Michiko Yuri, *Dynamical systems and ergodic theory*, vol. 40 of London Mathematical Society Student Texts, Cambridge University Press, Cambridge, UK, 1998.

[5] Pierre Tisseur, *Aspects ergodiques des automates cellulaires*, PhD thesis, Université d'Aix-Marseille II, 1999.