An inequality for the entropy of differentiable maps

*David Ruelle**

1. Introduction and statement of results.

The purpose of this note is to prove Theorem 2 below, which gives an upper bound to the measure-theoretic entropy $h(\rho)$ of any probability measure ρ invariant under a differentiable map f of a compact manifold M into itself. The upper bound is in terms of characteristic exponents introduced by the non-commutative ergodic theorem of Oseledec [2]. We first formulate a version of the latter theorem which will be suited to our purposes.

Theorem 1. *Let (M, Σ, ρ) be a probability space and $\tau : M \to M$ a measurable map preserving ρ. Let also $T: M \to \mathcal{M}_n(\mathbb{R})$ be a measurable map into the $m \times m$ matrices, such that**

$$\log^+ \|T(\cdot)\| \in L^1(M, \rho)$$

and write $$T_x^n = T(\tau^{n-1}x) \dots T(\tau x)\, T(x).$$

There is $\Omega \subset M$ such that $\rho(\Omega) = 1$ and for all $x \in \Omega$

$$\text{(1)} \qquad \lim_{n\to\infty} (T_x^{n*}\, T_x^n)^{1/2n} = \Lambda_x$$

exists [* denotes matrix transposition].

Let $\exp \lambda_x^{(1)} < \dots < \exp \lambda_x^{(s(x))}$ *be the eigenvalues of* Λ_x [with possibly $\lambda_x^{(1)} = -\infty$], *and* $U_x^{(1)}, \dots, U_x^{(s(x))}$ *the corresponding eigenspaces.* If $V_x^{(r)} = U_x^{(1)} + \dots + U_x^{(r)}$ *we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \|T_x^n u\| = \lambda_x^{(x)} \quad \text{when} \quad u \in V_x^{(r)} \backslash V_x^{(r-1)}$$

for $r = 1, \dots, s(x)$.

The theorem published by Oseledec assumes τ and T invertible. Its proof has been simplified by Raghunathan [4]. The above result can be obtained by modifying Raghunathan's argument.

Recebido em setembro de 1977.

* We write $\log^+ x = \max\{0, \log x\}$.

Let $m_x^{(r)} = \dim U_x^{(r)} = \dim V_x^{(r)} - \dim V_x^{(r-1)}$. The numbers $\lambda_x^{(1)}, \ldots, \lambda^{(s(x))}$, with multiplicities $m_x^{(1)}, \ldots, m_x^{(s(x))}$ constitute the *spectrum* of (ρ, τ, T) at x. The $\lambda_x^{(r)}$ are also called *characteristic exponents*. When n tends to ∞, $\frac{1}{n} \log \|T_x^n\|$ tends to the maximum characteristic exponent $\lambda_x^{(s(x))}$. The spectrum is τ-invariant; if ρ is τ-ergodic the spectrum is almost every where constant.

Let $T^{\wedge p} : M \to \mathcal{M}_{\binom{m}{p}}(\mathbb{R})$ be the p-th exterior power of T; we have

$$T^{\wedge p}(\tau^{n-1}x) \ldots T^{\wedge p}(\tau x)\, T^{\wedge p}(x) = (T_x^n)^{\wedge p}$$

and the spectrum of $(\rho, \tau, T^{\wedge p})$ is determined by

$$\lim_{n \to \infty} [(T_x^n)^{\wedge p *}(T_x^n)^{\wedge p}]^{\frac{1}{2n}} = \Lambda_x^{\wedge p}.$$

For $T^\wedge = \oplus_p T^{\wedge p}$ we obtain in particular

(2) $$\lim_{n\to\infty} \frac{1}{n} \log\ (T_x^n)^\wedge \ = \sum_{r:\lambda_x^{(r)}>0} \lambda_{m_x}^{(r)}\ \lambda_x^{(r)}$$

Theorem 2. *Let M be a C^∞ compact manifold and $f : M \to M$ a C^1 map Let I be the set of f-invariant probability measures on M:*

a) *There is a Borel subset Ω of M, such that $\rho(\Omega) = 1$ for every $\rho \in I$, and for each $x \in \Omega$ the following holds. There is a strictly increasing sequence of subspaces:*

$$0 = V_x^{(0)} \subset V_x^{(1)} \subset \ldots \subset V_x^{(s(x))} = T_x M$$

such that, for $r = 1, \ldots, s(x)$,

$$\lim_{n\to\infty} \frac{1}{n} \log \|T_x f^n u\| = \lambda_x^{(r)} \quad \text{if} \quad u \in V_x^{(r)} \backslash V_x^{(r-1)}$$

and $\lambda_x^{(1)} < \lambda_x^{(2)} < \ldots < \lambda_x^{(s(x))}$; we may have $\lambda_x^{(1)} = -\infty$. [The $V_x^{(r)}$ and $\lambda_x^{(r)}$ are uniquely defined with these properties, and independent of the choice of C^0 Riemann metric used to define $\|\cdot\|$]. The maps $x \to s(x)$, $(V_x^{(1)}, \ldots, V_x^{(s(x))})$, $(\lambda_x^{(1)}, \ldots, \lambda_x^{(s(x))})$ are Borel.

b) *Let $m_x^{(r)} = \dim V_x^{(r)} - \dim V_x^{(r-1)}$ for $r = 1, \ldots, s(x)$ and define*

$$\lambda_+(x) = \sum_{r:\lambda_x^{(r)}>0} m_x^{(r)}\, \lambda_x^{(r)}$$

Then, for every $\rho \in I$ the entropy $h(\rho)$ satisfies

$$h(\rho) \leq \rho(\lambda_+)$$

[*where* $\rho(\lambda_+) = \int \rho(dx)\lambda_+(x)$].

It is good to remember that the set I is convex and compact for the vague topology, and that $h : I \to \mathbb{R}$ is affine, but we shall not make use of these facts*.

We may assume that M has dimension m. Using a suitable Borel partition of M, we can trivialize the tangent bundle and write $TM \simeq M \times \mathbb{R}^m$. Therefore we can apply Theorem 1 with $\tau = f$, any $\rho \in I$, and $T(x)$ replaced by $T_x f$. We let Ω be the set of all x such that the limit (1) exists, and we take the $\lambda_x^{(r)}$ and $V_x^{(r)}$ as in Theorem 1. With these choices it is clear that part (*a*) of Theorem 2 holds. Part (*b*) is proved in Section 2.

2. Proof of the inequality $h(\rho) \leq \rho(\lambda_+)$.

In what follows we fix $\rho \in I$. We shall make use of the fact that, in view of (2),

(3) $$\lim_{n\to\infty} \frac{1}{n} \log \|(T_x f^n)^\wedge\| = \lambda_+(x).$$

Consider a smooth triangulation of M and for each m-dimensional simplex of the triangulation let there be a local chart such that the simplex is defined by

(4) $$t_1 \geq 0, \ldots, t_m \geq 0, \qquad t_1 + \ldots + t_m \leq 1.$$

It is convenient to assume that the boundary of each simplex has ρ-measure 0. This can be obtained by moving the triangulation by a small diffeomorphism of M (one pushes the triangulation successively by vector fields with small compact supports covering M so that the mass of the boundaries becomes zero). Given an integer $N > 0$, we decompose the simplex (4) into subsets by the planes

$$t_1 = \frac{k_1}{N} \quad \text{for} \quad i = 1, \ldots, N-1.$$

We can assume that these planes have ρ-measure 0 for all N (use a small diffeomorphism of the simplex reducing to the identity on the boundary).

*They could be used to reduce the proof of the inequality $h(\rho) \leq \rho(\lambda_+)$ to the case where ρ is ergodic.

We have thus obtained a partition δ_N of M (up to sets of measure zero) into cubes and (near the boundary of the simplexes) pieces of cubes.

a) *Given a Riemann metric on M, there is $C > 0$, and for each n there is $N(n)$, such that if $N > N(n)$ the number of sets of δ_N intersected $f^n S$ where $S \in \delta_N$ is less than*

$$C \| T_x f^n)^\wedge \| \tag{5}$$

for any $x \in S$.

Since N is large, diam S is small, and f^n restricted to S is close to its linear part estimated at any $x \in S$ when computed in terms of the variables t_i corresponding to the simplex in which S lies and to the simplex(es) in which $f^n S$ lies. Using the equivalence of the Riemann metric on M and of the Euclidean metric in the variables t_i, we find that there is $K > 0$ (independent of n, N) such that $f^n S$ lies in a rectangular parallelepiped with sides $K\frac{a_1}{N}, \dots, K\frac{a_p}{N}, \frac{K}{N}, \dots, \frac{K}{N}$, where $a_1, \dots, a_p > 1$ and

$$a_1, \dots, a_p = \max\{\|(T_x f^n)^\wedge u\| : u \in (T_x M)^\wedge, \|u\| = 1\} = T_x f^n)^\wedge\|.$$

Now, a cube of sides $\frac{1}{N}$ can intersect only a bounded number of sets in the decomposition of a simplex by planes $t_i = \frac{k_i}{N}\cdot$ Therefore the number of sets of δ_N intersected by $f^n S$ is bounded by an expression of the form (5).

b) *The entropy of ρ with respect to f^n and the partition δ_N satisfies*

$$h_{f^n}(\rho, \delta_N) \le \log C + \int \rho(dx) \log \| (T_x f^n)^\wedge \| . \tag{6}$$

Each $x \in M$ is in some $S = S_0 \cap S_1 \cap \dots \cap S_{k-1}$ where $S_j \in f^{-nj}\delta_N$, and we can define

$$h_{N,n,k}(x) = - \sum_{S_k \in f^{-nk}\delta_N} \frac{\rho(S \cap S_k)}{\rho(S)} \log \frac{\rho(S \cap S_k)}{\rho(S)}$$

Then

$$h_{f^n}(\rho, \delta_N) = \lim_{\ell \to \infty} \frac{1}{\ell} \sum_{k=0}^{\ell-1} \int \rho(dx) h_{N,n,k}(x)$$

and, for $k > 0$, (a) yields

$$h_{N,n,k}(x) \le \log [C \|(T_x f^n)^\wedge\|].$$

Therefore (6) holds.

c) *End of proof.*
Letting N tend to $+\infty$ in (6) and dividing by n we obtain

$$h_f(\rho) = \frac{1}{n} h_{f^n}(\rho) \le \frac{1}{n} \log C + \int \rho(dx) \frac{1}{n} \log \|(T_x f^n)^\wedge\|.$$

Since $\frac{1}{n} \log \|(T_x f^n)^\wedge\|$ is positive and bounded above, (3) permits to conclude that

$$h_f(\rho) \le \int \rho(dx)\, \lambda_+(x).$$

3. Remark

The inequality $h(\rho) \le \rho(\lambda_+)$ was known for axiom A diffeomorphisms and for the time one map of axiom A flows [5], [6]. It is also obvious for quasi-periodic maps of the m-torus. A related result was proved for certain diffeomorphisms preserving a smooth measure by Margulis and Pesin [3]. In all those cases one has

$$\sup_\rho [h(\rho) - \rho(\lambda_+)] = 0.$$

Question. Is this "variational principle" true in general?

References

[1] R. Bowen and D. Ruelle, *The ergodic theory of axiom A flows*. Inventiones math. *29*, 181-202 (1975).

[2] V. I. Oseledec, *A multiplicative ergodic theorem. Ljapunov characteristic numbers for dynamical systems*. Trudy Moskov. Mat. Obsc. *19*, 179-210 (1968). English. Translation. Trans. Moscow Math. Soc. *19*, 197-231 (1968).

[3] Ja. B. Pesin, *Ljapunov characteristic exponents and ergodic properties of smooth dynamical systems with an invariant measure*. Dokl. Akad. Naul. SSSR *226* N.° 4, 774-777 (1976). English translation. Soviet. Math. Dokl. *17* N.° 1, 196-199 (1976).

[4] M. S. Raghunathan, *A proof of Oseledec' multiplicative ergodic theorem*. Unpublished.

[5] D. Ruelle, *A measure associated with axiom A atractors*. Amer. J. Math. *98*, 619-654 (1976).

Instituto de Matemática Pura e Aplicada,
Rio de Janeiro, Brasil

Institut des Hautes Études Scientifiques,
Bures-sur-Yvette, França