Astérisque

MICHAL MISIUREWICZ

# A short proof of the variational principle for a $\mathbb{Z}_+^N$ action on a compact space

*Astérisque*, tome 40 (1976), p. 147-157
<http://www.numdam.org/item?id=AST_1976__40__147_0>

© Société mathématique de France, 1976, tous droits réservés.

L'accès aux archives de la collection « Astérisque » (http://smf4.emath.fr/Publications/Asterisque/) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
*Article numérisé dans le cadre du programme Numérisation de documents anciens mathématiques*
http://www.numdam.org/

Société Mathématique de France
Astérisque 40 (1976) p.147-157

# A SHORT PROOF OF THE VARIATIONAL PRINCIPLE FOR A $\mathbb{Z}_+^N$ ACTION ON A COMPACT SPACE

Michał Misiurewicz

0. Introduction. Ruelle in [12] introduced the notion of pressure for an action of the group $\mathbb{Z}^N$ on a compact metric space. It is a generalization of the notion of topological entropy. The variational principle (proved in [12] under some strong conditions) is a generalization of the Dinaburg's theorem ([5,8,7]) on a connection between the topological and measure entropies. A general proof of the variational principle was given by Walters [13] (see also Denker [3]) for an action of $\mathbb{Z}_+$ and by Elsanousi in [6] for an action of $\mathbb{Z}_+^N$.

The first part of the proof given below (an inequality $h_\mu(T)+\mu f \leq P(T,f)$ ) is a natural generalization of the proof from [10]. The second part ( $\sup_\mu (h_\mu(T) + \mu f) \geq P(T,f)$ ) is quite new (although the idea is close to the Ruelle's one).

1. Notations.

$\mathbb{Z}_+$ denotes the set of all non-negative integers. Let us fix a positive integer $N$ .

$G = \mathbb{Z}_+^N$ is a commutative semigroup with respect to addition. For $n \in G$ we denote by $n_i$ the i-th coordinate of $n$ $(i=1,\dots,N)$ . For $n,m \in G$ let $nm = (n_1m_1,\dots,n_Nm_N)$ . The relation $\geq$ ($n \geq m$ iff $n_i \geq m_i$ for $i=1,\dots,N$) directs $G$ .

X is a non-empty compact Hausdorff space.

C(X) is the space of all continuous real functions on X with the norm $\|f\| = \sup_{x \in X} |f(x)|$ .

$\mathcal{M}(X)$ is the space of all positive Borel regular normed measures on $X$ . It can be identified with the space of all positive linear functionals on $C(X)$ having norm 1 (therefore we shall write $\mu f$ instead of $\int_X f\, d\mu$ for $f \in C(X)$ ). We consider the weak-$^*$ topology on $\mathcal{M}(X)$ (then $\mathcal{M}(X)$ is compact).

$T$ $(n \longmapsto T^n)$ is an action of $G$ on $X$ , i.e. a homomorphism of $G$ into the semigroup of all continuous transformations of $X$ into itself (i.e. $T^n : X \longrightarrow X$ , $T^{n+m} = T^n \circ T^m$ ).

$T^{*n} : C(X) \longrightarrow C(X)$ (for $n \in G$) is the operator induced by $T^n$ (i.e. $T^{*n} f = f \circ T^n$ ).

$T^{**n} : \mathcal{M}(X) \longrightarrow \mathcal{M}(X)$ (for $n \in G$) is a restriction of the operator induced by $T^{*n}$ (i.e. $T^{**n}\mu = \mu \circ T^{*n}$ ) to $\mathcal{M}(X)$ . It is easy to check that indeed $T^{**n}(\mathcal{M}(X)) \subset \mathcal{M}(X)$ and that $T^{**n}$ is continuous.

Of course $T^*$ and $T^{**}$ are actions of $G$ on $C(X)$ and $\mathcal{M}(X)$ , respectively.

$\mathcal{M}(X,T)$ is a space of all T-invariant measures (i.e. these elements of $\mathcal{M}(X)$ which are fixed points of all $T^{**n}$ , $n \in G$ ).

$\mathcal{W}$ is the set of all neighbourhoods of the diagonal in $X \times X$ , directed by the inclusion. It is a uniform structure (uniformity) for $X$ (see [9] ).

## 2. Definitions of pressure and entropy.

We may in the natural way extend Ruelle's definition of pressure of $\mathbb{Z}^N$ -action ([12] ).

Let $n \in G$ , $\delta \in \mathcal{W}$, $f \in C(X)$ . We define successively:

$$\Lambda(n) = \{m \in G : m_i < n_i \quad \text{for} \quad i = 1, \ldots, N\} ,$$

$$\lambda(n) = \operatorname{Card} \Lambda(n) = n_1 \cdot \ldots \cdot n_N ,$$

$$\delta_n = \bigcap_{k \in \Lambda(n)} (T^k \times T^k)^{-1} \delta ,$$

$$f_n = \sum_{k \in \Lambda(n)} T^{*k} f .$$

Of course $\delta_n \in \mathscr{H}$, $f_n \in C(X)$ .

A finite subset $e$ of $X$ is called:

$(n,\delta)$-separated, if for any $x,y \in e$ , $x \neq y$ , we have $(x,y) \notin \delta_n$ ,

$(n,\delta)$-spanning, if for any $x \in X$ there exists $y \in e$ such that $(x,y) \in \delta_n$ .

We denote $p(f,e) = \log \sum_{x \in e} \exp f(x)$ . We define further

$$(1) \qquad P_{n,\delta}(T,f) = \sup\{p(f_n,e) : e \text{ is } (n,\delta)\text{-separated}\}$$

$$(2) \qquad P_\delta(T,f) = \limsup_{n \in G} \frac{1}{\lambda(n)} P_{n,\delta}(T,f)$$

Of course, $P_\delta(T,f) \geq P_\varepsilon(T,f)$ for $\delta \subset \varepsilon$ . Therefore it is possible to define the pressure

$$(3) \qquad P(T,f) = \lim_{\delta \in \mathscr{H}} P_\delta(T,f) = \sup_{\delta \in \mathscr{H}} P_\delta(T,f)$$

In the sequel we shall use the measure entropy function $H_\mu(\cdot)$ (for the definition and properties see e.g. [11]) .

For $\mu \in \mathscr{M}(X,T)$ the entropy $h_\mu(T)$ may be defined in the same way as in [2] for the action of $\mathbb{Z}^N$ . For a Borel finite partition $A$ of the space $X$ we define

$$A^n = \bigvee_{k \in \Lambda(n)} (T^k)^{-1}A \quad \text{for} \quad n \in G ,$$

$h_\mu(T,A) = \lim_{n \in G} \frac{1}{\lambda(n)} H_\mu(A^n)$ , (it is easy to show that the limit exists, cf. [2]).

Finally,

$$h_\mu(T) = \sup\{h_\mu(T,A) : A \text{ - Borel finite partition}\} .$$

## 3. The variational principle.

We shall prove the following variational principle:

$P(T,f) = \sup_{\mu \in \mathscr{M}(X,T)} (h_\mu(T) + \mu f)$ . The proof consists of two parts.

Part I.

$$h_\mu(T) + \mu f \le P(T,f) \quad \text{for} \quad \mu\in\mathcal{M}(X,T) .$$

Proof.

Let $\mu\in\mathcal{M}(X,T)$ . Let us fix $\xi > 0$ and a Borel finite partition $A$ . Take $m \in G$ such that

$$(4) \qquad \log 2 \le \xi \cdot \lambda(m)$$

Let $A^m$ consist of the sets $a_1,\dots,a_s$ . For any of them there exists a compact set $b_i \subset a_i$ such that

$$\mu(a_i \setminus b_i) \le \frac{\xi}{s \log s} .$$

Let $b_o = X \setminus \bigcup_{i=1}^{s} b_i$ . For the partition $B = \{b_o,b_1,\dots,b_s\}$ we have

$$(5) \qquad H_\mu(A^m|B) \le \mu(b_o)\cdot \log s \le \xi .$$

We take

$$\varepsilon = (X \times X) \setminus \bigcup_{\substack{i,j=1\\ i\ne j}}^{s} (b_i \times b_j) \in \mathcal{U}$$

and next $\delta\in\mathcal{U}$ such that $\delta\circ\delta\subset\varepsilon$ (i.e. if $(x,y),(y,z)\in\delta$ , then $(x,z)\in\varepsilon$ ) and $|f(x) - f(y)| \le \xi$ if $(x,y)\in\delta$ .

Let us fix $n \in G$ . There exists a maximal $(nm,\delta)$-separated (i.e. being also $(nm,\delta)$-spanning) set $e\subset X$ .

Denote by $C = \bigvee_{k\in\Lambda(n)} (T^{km})^{-1}B$ .

Further, denote

$$\alpha(b) = \sup_{x\in b} f_{nm}(x) \quad \text{for} \quad b\in C \;, \quad \beta = \sum_{b\in C} \exp \alpha(b) .$$

We have

$$\int_b f_{nm}\, d\mu \le \alpha(b)\cdot\mu(b) ,$$

therefore

$$H_\mu(C) + \mu f_{nm} \le \sum_{b\in C} \mu(b)(\alpha(b) - \log\mu(b)) = \beta\cdot \sum_{b\in C} \frac{\exp\alpha(b)}{\beta}\cdot\eta\left(\frac{\mu(b)}{\exp\alpha(b)}\right) ,$$

where $\eta(x) = -x \log x$ .

The function $\eta$ is concave, therefore

$$(6) \qquad H_\mu(C) + \mu f_{nm} \leq \beta \cdot \eta\Big(\sum_{b \in C} \frac{\exp\alpha(b)}{\beta} \cdot \frac{\mu(b)}{\exp\alpha(b)}\Big) = \log \beta$$

e is a $(nm,\delta)$-spanning set, hence for every $b \in C$ there exists a point $z(b) \in e$ such that

$$\alpha(b) = \sup\{f_{nm}(x) : x \in b , (x,z(b)) \in \delta_{nm}\} .$$

But if

$$(x,z(b)) \in \delta_{nm} ,$$

then for $k \in \Lambda(nm)$

$$(T^k x, T^k z(b)) \in \delta ,$$

thus, in view of the definition of $\delta$ ,

$$|(T^{*k} f)(x) - (T^{*k} f)(z(b))| \leq \xi .$$

Hence

$$(7) \qquad f_{nm}(z(b)) \geq \alpha(b) - \xi \cdot \lambda(nm)$$

From the definitions of $\delta$ and $\varepsilon$ we obtain for $y \in e$ , $k \in \Lambda(n)$ : $\text{Card}\,\{a \in B : \exists_{x \in a} (T^{km}x, T^{km}y) \in \delta\} \leq 2$ , thus for $y \in e$ $\text{Card}\,\{b \in C : \exists_{x \in b} (x,y) \in \delta_{nm}\} \leq 2^{\lambda(n)}$ ,
because

$$\delta_{nm} \subset \bigcap_{k \in \Lambda(n)} (T^{km} \times T^{km})^{-1} \delta .$$

Hence

$$(8) \qquad \text{Card}\,\{b \in C : z(b) = y\} \leq 2^{\lambda(n)}$$

Hence, from (7) and (8) we get:

$$2^{\lambda(n)} \cdot \sum_{y \in e} \exp f_{nm}(y) \geq \sum_{b \in C} \exp \alpha(b) \cdot \exp(-\xi \cdot \lambda(nm)) ,$$

thus,

$$(9) \qquad \lambda(n)\cdot\log 2 + p(f_{nm},e) \geq \log\beta - \xi\cdot\lambda(nm)$$

But $\mu f_{nm} = \lambda(nm)\cdot\mu f$ , so from this, from (1),(4),(6) and (9) we obtain (notice that $\lambda(nm) = \lambda(n)\cdot\lambda(m)$ ) :

$$(10) \qquad \frac{1}{\lambda(nm)} H_\mu(C) + \mu f \leq \frac{1}{\lambda(nm)} P_{nm,\delta}(T,f) + 2\xi$$

In view of (5), for $k\in\Lambda(n)$ we have

$$H_\mu((T^{km})^{-1}A^m | (T^{km})^{-1}B) \leq \xi ,$$

therefore

$$H_\mu(A^{nm}|C) = H_\mu(\bigvee_{k\in\Lambda(n)} (T^{km})^{-1}A^m | \bigvee_{k\in\Lambda(n)} (T^{km})^{-1}B) \leq \xi\cdot\lambda(n) .$$

Hence

$$H_\mu(A^{nm}) \leq H_\mu(C) + H_\mu(A^{nm}|C) \leq H_\mu(C) + \xi\cdot\lambda(n) ,$$

thus we obtain from (10):

$$\frac{1}{\lambda(nm)} H_\mu(A^{nm}) + \mu f \leq \frac{1}{\mu(nm)} P_{nm,\delta}(T,f) + 3\xi .$$

Taking lim sup with respect to n we get

$$h_\mu(T,A) + \mu f \leq P_\delta(T,f) + 3\xi \leq P(T,f) + 3\xi .$$

But $\xi$ and A were arbitrary, hence

$$h_\mu(T) + \mu f \leq P(T,f) .$$

Part II.

$$\sup_{\mu\in\mathcal{M}(X,T)} (h_\mu(T) + \mu f) \geq P(T,f) .$$

Proof.

Let us fix $\delta\in\mathcal{H}$ . For every $n\in G$ we choose such an $(n,\delta)$-separated set $e_n$ that

$$(11) \qquad p(f_n,e_n) \geq P_{n,\delta}(T,f) - 1$$

Let us define a measure $\sigma_n$, concentrated on $e_n$, by a formula

$$\sigma_n(\{y\}) = \exp(f_n(y) - p(f_n, e_n)) \text{ for } y \in e_n .$$

We have

$$\sum_{y \in e_n} \sigma_n(\{y\}) = 1 ,$$

therefore

$$\sigma_n \in \mathcal{M}(X) .$$

Let

$$\mu_n = \frac{1}{\lambda(n)} \sum_{k \in \Lambda(n)} T^{**k} \sigma_n .$$

For some sequence $(n_i)_{i=1}^{\infty}$ cofinal with $G$ we have

$$(12) \qquad \lim_{i \to \infty} \frac{1}{\lambda(n_i)} P_{n_i, \delta}(T, f) = P_\delta(T, f)$$

We choose some cluster point of the sequence $(\mu_{n_i})_{i=1}^{\infty}$ and denote it by $\mu$. Of course, $\mu$ is also a cluster point of the net $(\mu_n)_{n \in G}$. For $g \in C(X)$ and $k \in G$ fixed, the function

$$\Phi : \mathcal{M}(X) \longrightarrow \mathbb{R} ,$$

given by the formula

$$\Phi\nu = \nu g - \nu(T^{*k} g) ,$$

is continuous, therefore $\Phi\mu$ is a cluster point of the net $(\Phi\mu_n)_{n \in G}$. We have

$$|\Phi\mu_n| \leq \frac{1}{\lambda(n)} \cdot 2 \cdot (\lambda(n) - \lambda(n-k)) \cdot \|g\| \quad (\text{for } n \geq k) ,$$

because

$$\text{Card}\,(\Lambda(n) \setminus (k + \Lambda(n))) = \text{Card}\,((k + \Lambda(n)) \setminus \Lambda(n)) = \lambda(n) - \lambda(n-k) .$$

But

$$\lim_n \frac{\lambda(n-k)}{\lambda(n)} = 1 ,$$

thus

$$\Phi\mu = 0 .$$

Hence $\mu g = (T^{**k}\mu) g$. But $g$ and $k$ were arbitrary, therefore

$$\mu \in \mathcal{M}(X, T) .$$

There exists a Borel finite partition $A$ of $X$ such that

$a \times a \subset \delta$ for $a \in A$. Then for $a \in A^n$ $a \times a \subset \delta_n$, therefore

$$\mathrm{Card}\,(e_n \cap a) \leq 1 .$$

Hence

$$H_{\sigma_n}(A^n) + \sigma_n f_n = \sum_{y \in e_n} \sigma_n(\{y\})(f_n(y) - \log \sigma_n(\{y\})) = p(f_n, e_n) .$$

Let us fix $m, n \in G$, $n \geq 2m$. For given $j \in \Lambda(m)$ let

$$s(j) = \left(E\left(\frac{n_1 - j_1}{m_1}\right), \dots, E\left(\frac{n_N - j_N}{m_N}\right)\right) .$$

We have:

$$A^n = \bigvee_{r \in \Lambda(s(j))} (T^{rm+j})^{-1} A^m \vee \bigvee_{k \in \Xi} (T^k)^{-1} A ,$$

where

$$\Xi = \Lambda(n) \setminus (j + \Lambda(ms(j))) .$$

But

$$\mathrm{Card}\,\Xi = \lambda(n) - \lambda(ms(j)) \leq \lambda(n) - \lambda(n-2m) ,$$

thus

$$p(f_n, e_n) = H_{\sigma_n}(A^n) + \sigma_n f_n \leq$$

$$\leq \sum_{r \in \Lambda(s(j))} H_{\sigma_n}((T^{rm+j})^{-1} A^m) + \sigma_n f_n + (\lambda(n) - \lambda(n-2m)) \log \mathrm{Card}\ A.$$

Summing the inequalities obtained for $j \in \Lambda(m)$ we get (notice that for $k \in \Lambda(n)$ there exists a unique $j \in \Lambda(m)$ and a unique $r \in \Lambda(s(j))$ such that $k = rm + j$ ) :

$$(13) \qquad \sum_{k \in \Lambda(n)} H_{\sigma_n}((T^k)^{-1} A^m) + \lambda(m) \cdot \sigma_n f_n \geq$$

$$\geq \lambda(m) \cdot (p(f_n, e_n) - (\lambda(n) - \lambda(n-2m)) \cdot \log \mathrm{Card}\ A)$$

We have also

$$(14) \qquad \sigma_n f_n = \sigma_n\Big(\sum_{k \in \Lambda(n)} T^{*k} f\Big) = \Big(\sum_{k \in \Lambda(n)} T^{**k} \sigma_n\Big) f = \lambda(n) \cdot \mu_n f$$

From the definition of entropy it follows that

$$(15) \qquad H_{\sigma_n}((T^k)^{-1}A^m) = H_{T^{**k}\sigma_n}(A^m)$$

From the definition of entropy and from the concavity of the function $-x \log x$ it follows that

$$(16) \qquad H_{\mu_n}(A^m) \geq \frac{1}{\lambda(n)} \sum_{k\in\Lambda(n)} H_{T^{**k}\sigma_n}(A^m)$$

The formulas (11) and (13) - (16) give us

$$(17) \qquad \frac{1}{\lambda(m)} H_{\mu_n}(A^m) + \mu_n f \geq \frac{1}{\lambda(n)} P_{n,\delta}(T,f) - \frac{1}{\lambda(n)} \cdot$$

$$\cdot((\lambda(n) - \lambda(n-2m)) \log \operatorname{Card} A + 1)$$

The partition $A$ can be chosen in such a way that the boundaries of elements of $A$ have measure $\mu$ zero. (see [1], Chapt.IV, §5, exerc. 13 d; see also [10]) . Then $A^n$ has the same property. But for a set $a$ with the boundary of measure $\mu$ zero, the function $\mathcal{M}(X)\longrightarrow\mathbb{R}$ , given by $\nu\longmapsto\nu(a)$ , is continuous in the point $\mu$ . Hence the function $\mathcal{M}(X)\longrightarrow\mathbb{R}$ , given by $\nu\longmapsto H_\nu(A^m)$ is also continuous at the point $\mu$ , therefore in view of (12), (17) and the definition of $\mu$ , we have

$$\frac{1}{\lambda(m)} H_\mu(A^m) + \mu f \geq P_\delta(T,f) .$$

Taking the limit with respect to $m$ and using the inequality $h_\mu(T,A) \leq h_\mu(T)$ , we obtain

$$h_\mu(T) + \mu f \geq P_\delta(T,f)$$

But $\delta$ was arbitrary, hence

$$\sup_{\nu\in\mathcal{M}(X,T)} (h_\nu(T) + \nu f) \geq P(T,f) .$$

4. Remark..

If $h_\mu(T) + \mu f = P(T,f)$ , then $\mu$ is called an equilibrium state for $(T,f)$ (measure with maximal entropy in the case of $f = 0$). The above construction shows that if $P_\delta(T,f) = P(T,f)$ for some $\delta \in \mathscr{W}$, then there exists an equilibrium state for $(T,f)$ . In the case of $N = 1$ , $f = 0$ , this can be reformulated as follows:

If there exists an open cover $A$ such that $h(T,A) = h(T)$ , then there exists a measure with maximal entropy.

This is a particular case of the theorem of Denker ([4]) , but obtained without assuming $X$ finite dimensional.

REFERENCES


[1] N. Bourbaki, Éléments de Mathématique, Livre VI (Intégration) Chap. IV - Hermann & C$^{ie}$ , Paris, 1952.

[2] J.P. Conze, Entropie d'un groupe abélien de transformations, Z. Wahrscheinlichkeitstheorie verw. Geb., 25 (1973), 11 - 30.

[3] M. Denker, Remarques sur la pression pour les transformations continues, C.R. Acad. Sci. Paris, 279 (1974), 967 - 970.

[4] M. Denker, Measures with maximal entropy, preprint

[5] E.I. Dinaburg, Связь между различными энтропийными характеристиками динамических систем - Изв. А. Н. СССР, 35 (1971), 324 - 366.

[6] S.A. Elsanousi, A variational principle for the pressure of a continuous $\mathbb{Z}^2$-action on a compact metric space, preprint.

[7] T.N.T. Goodman, Relating topological entropy and measure entropy, Bull. Lond. Math. Soc., 3 (1971), 176 - 180 .

[8] L.W. Goodwyn, Topological entropy bounds measure-theoretic entropy, Proc. Amer. Math. Soc., 23 (1969), 679 - 688.

[9] J.L. Kelley, General Topology, Van Nostrand, Princeton, 1955.

[10] M. Misiurewicz, Topological conditional entropy, Studia Math., 55 (1976), 175 - 200.

[11] V.A. Rohlin, Лекции по энтропийной теории преобразований с инвариантной мерой - УМН, 22 (1967), 3 - 56. (Lectures on the entropy theory of measure preserving transformations, Russian Math. Surv. 22,(1967), 1 - 52)

[12] D. Ruelle, Statistical mechanics on a compact set with $\mathbb{Z}^\nu$ action satisfying expansiveness and specification, Trans. Amer. Math. Soc., 185 (1973), 237 - 252.

[13] P. Walters, A variational principle for the pressure of continuous transformations. preprint.

Mathematical Institute, Univ. of Warsaw
PKiN IX p.
00-901 Warszawa
Poland.