ANNALES DE L'I. H. P., SECTION A

JEAN BRICMONT

On the transport of equilibrium states

Annales de l'I. H. P., section A, tome 28, nº 4 (1978), p. 423-429
<http://www.numdam.org/item?id=AIHPA_1978__28_4_423_0>

Ann. Inst. Henri Poincaré,
Vol. XXVIII, n° 4, 1978, p. 423-429.

Section A :
Physique théorique.

On the transport of equilibrium states

by

Jean BRICMONT †

Institut de Physique Théorique,
Université Catholique de Louvain

ABSTRACT. — We study the relation between the equilibrium states of continuous transformations on compact metric spaces connected by a continuous surjection.

RÉSUMÉ. — On étudie la relation entre les états d'équilibre de transformations continues sur des espaces métriques compacts reliés par une surjection continue.

1. INTRODUCTION

The purpose of this note is to develop an argument of Bowen [*1*] concerning the transport of the topological entropy under a surjective map to the transport of the topological pressure and then of the equilibrium states. Namely assume that there is a continuous surjection from one compact metric space X onto another Y; let A be a continuous function from Y to $\mathbb{R}$. If the inverse images under f of points in Y satisfy some homogeneity condition then the map f induces a surjection from the equilibrium states of $A \circ f$ onto the equilibrium states of A.

Our result is in fact an application of a general form of the variational principle due to Ledrappier and Walters [*7*].

In this section, we develop the necessary formalism and in the next one we state and prove the theorem. The last section is devoted to examples. The framework applies quite generally to groups with endomorphisms and, as a special case, to classical lattice spin systems.

Let (X, d) be a compact metric space. $\mathscr{C}(X, X)$ is the set of continuous functions from X to X. $\mathbb{Z}_+$ is the set of non-negative integers.

Let T be a map from $\mathbb{Z}_+^\nu$ to $\mathscr{C}(X, X)$ such that $T(i + j) = T(i) \circ T(j)$. $T(i)$ will be noted T^i.

Given $a \in \mathbb{Z}_+^\nu$, we denote by

$$\Lambda(a) = \{ x = (x_1 \ldots x_\nu) \in \mathbb{Z}_+^\nu) x_i \leq a_i \quad i = 1 \ldots \nu \}$$

and by $a \to \infty$ we mean $a_i \to \infty$, $i = 1 \ldots \nu$.

DÉFINITION. — Let $\varepsilon > 0$, $\Lambda \in \mathscr{P}_f(\mathbb{Z}_+^\nu)$ and K be a compact subset of X.

A subset E of K is (Λ, ε)-*separated* if $x, y \in E$ and $\max_{i\in\Lambda} d(T^i x, T^i y) \leq \varepsilon$ implies $x = y$.

A subset F of K is (Λ, ε)-*spanning* in K if $\forall x \in K$, $\exists y \in F$ such that

$$\max_{i\in\Lambda} d(T^i x, T^i y) \leq \varepsilon$$

Let

$$Z_\Lambda^{(1)}(A, \varepsilon, T, K) = \sup \left\{ \sum_{x\in E} \exp \sum_{i\in\Lambda} A(T^i x) / E \text{ is } (\Lambda, \varepsilon)\text{-separated and } E \subset K \right\}$$

$$Z_\Lambda^{(2)}(A, \varepsilon, T, K) = \inf \left\{ \sum_{x\in F} \exp \sum_{i\in\Lambda} A(T^i x) / F \text{ is } (\Lambda, \varepsilon)\text{-spanning and } F \subset K \right\}$$

THEOREM 1 [2] [3]. — Let

$$P^{(i)}(A, \varepsilon, T, K) = \lim_{a\to\infty} \sup \frac{1}{|\Lambda(a)|} \log Z_{\Lambda(a)}^{(i)}(A, \varepsilon, T, K) \qquad i = 1, 2$$

Then $\lim_{\varepsilon\to 0} P^{(i)}(A, \varepsilon, T, K) = P(A, T, K)$ exists (possibly $= +\infty$).

If $K = X$ we write $P(A, T, X) = (P(A, T)$.

— $P(A, T)$ is called the *topological pressure of* A

— $P(0, T)$ is the *topological entropy*

— $P(A, T)$ is finite for all $A \in \mathscr{C}(X, \mathbb{R})$ iff $P(0, T)$ is finite

— If $P(0, T)$ is finite, $P(., T)$ is a convex continuous function on $\mathscr{C}(X, \mathbb{R})$

Let $\mathscr{M}(X)$ be the set of probability measures on X; I(X) *is the set of* $\mathbb{Z}_+^\nu$ *invariant elements* of $\mathscr{M}(X)$:

$$\mu \in I(X) \quad \text{if} \quad \mu \in \mathscr{M}(X)$$

and

$$\mu(A \circ T^i) = \mu(A) \qquad \forall A \in \mathscr{C}(X, \mathbb{R}), \quad \forall i \in \mathbb{Z}_+^\nu$$

Let $h_\mu(T)$ denote the measure-theoretic entropy of T with respect to μ, then the following variational principle holds [2] [3]:

$$P(A, T) = \sup \{ h_\mu(T) + \mu(A) \mid \mu \in I(X) \} \tag{1}$$

More generally, if Y is a compact metric space with a $\mathbb{Z}_+^\nu$ action denoted by S and f is a continuous surjection from X onto Y with $f \circ T^i = S^i \circ f$, then [7],

$$\forall A \in \mathscr{C}(X, \mathbb{R}), \quad \forall \nu \in I(Y)$$
$$\sup \{ h_\mu(T) + \mu(A) \mid f^*(\mu) = \nu \} = h_\nu(S) + \int P(A, T, f^{-1}(y)) d\nu(y) \quad (2)$$

where f^* is the map from $\mathscr{M}(X)$ onto $\mathscr{M}(Y)$ induced by f:

$$f^*(\mu)(A) = \mu(A \circ f), \qquad \forall A \in \mathscr{C}(Y, \mathbb{R})$$

We introduce the set I_A of *equilibrium states of* A:

$$I_A = \{ \mu \in I(X) \mid h_\mu(T) + \mu(A) = P(A, T) \}$$

If P(0, T) is finite, then [5]

$$I_A = \{ \mu \in I(X) \mid P(A + B, T) - P(A, T) \geq \mu(B), \qquad \forall B \in \mathscr{C}(X, \mathbb{R}) \} \neq \emptyset$$

2. THE RESULT

We introduce now the assumptions that we shall use; in the next section we will give examples of this structure.

(A.1) (X, d), (Y, e), are compact metric spaces.

(A.2) T(resp. S) is a map from $\mathbb{Z}_+^\nu$ to $\mathscr{C}(X, X)$(resp. $\mathscr{C}(Y, Y)$) such that $T^{i+j} = T^i \circ T^j$ (resp. $S^{i+j} = S^i \circ S^j$).

(A.3) *i)* $f : X \to Y$ is a continuous and surjective map
ii) $\forall i \in \mathbb{Z}_+^\nu$, $f \circ T^i = S^i \circ f$

(A.4) $P(0, T, f^{-1}(y))$ is constant for $y \in Y$.

(A.5) *i)* $(G, \tilde{d})$ is a compact metric space
ii) R is a map from $\mathbb{Z}_+^\nu$ to $\mathscr{C}(G, G)$ and $R(i + j) = R(i) \circ R(j)$
iii) $H : X \times G \to X$ is a continuous map such that $H(x, g) = H(x, g')$ for some $x \in X$, implies $g = g'$. $H(x, g)$ is denoted by xg.
iv) $f^{-1}(f(x)) = xG$, $\forall x \in X$
v) $\forall i \in \mathbb{Z}_+^\nu$, $\forall x \in X$, $\forall g \in G$, $T^i(xg) = T^i(x)R^i(g)$.

The main results are:

THEOREM 2. — Under the assumptions (A.1)-(A.4) above and if P(0, T) is finite for any $A \in \mathscr{C}(Y, \mathbb{R}) f^*$ is a surjection from $I_{A \circ f}$ onto I_A.

LEMMA. — Under the assumptions of Theorem 2,

$$P(A \circ f, T) = P(A, S) + P(0, T, f^{-1}(y))$$

THEOREM 3. — Under the assumptions (A.1), (A.2), (A.3), (A.5), assumption (A.4) is satisfied and the conclusion of Theorem 2 holds; in fact,

$$\forall y \in Y, \qquad P(0, T, f^{-1}(y)) = P(0, R, G)$$

Proof of Theorem 2. — Let $\mu \in I_{A\circ f}$; then for all $B \in \mathscr{C}(Y, \mathbb{R})$

$$P(A \circ f + B \circ f, T) - P(A \circ f, T) \geqslant \mu(B \circ f)$$

By the definition of f^*, $\mu(B \circ f) = f^*(\mu)(B)$.

Therefore, by the Lemma: $\forall B \in \mathscr{C}(Y, \mathbb{R})$

$$f^*(\mu)(B) \leqslant P(A + B, S) - P(A, S)\,;$$

this means that $f^*(\mu) \in I_A$. (By Assumption (A.3 *ii*), $f^*(\mu)$ is $\mathbb{Z}_+^\nu$ invariant if μ is).

Given $\nu \in I_A$, we will show that there exists a $\mu \in I_{A\circ f}$ such that $f^*(\mu) = \nu$. For those elements of $\mathscr{C}(X, \mathbb{R})$ of the form $A \circ f$ with $A \in \mathscr{C}(Y, \mathbb{R})$, we define $\bar{\mu}(A \circ f) = \nu(A)$.

Since $\nu \in I_A$, $\nu(B) \leqslant P(A + B, S) - P(A, S)$, therefore, by the lemma,

$$\bar{\mu}(B \circ f) \leqslant P(A \circ f + B \circ f, T) - P(A \circ f, T).$$

Since P is convex, using the Hahn-Banach Theorem, $\bar{\mu}$ can be extended to a measure μ on $\mathscr{C}(X, \mathbb{R})$ satisfying, for all $C \in \mathscr{C}(X, \mathbb{R})$

$$\mu(C) \leqslant P(A \circ f + C, T) - P(A \circ f, T)$$

One can see from the Definition of the pressure that

$$P(A \pm (B \circ T^i - B), T) - P(A, T) = 0, \qquad \forall A, B \in \mathscr{C}(X, \mathbb{R}), \quad \forall i \in \mathbb{Z}_+^\nu$$

Therefore $\mu(C \circ T^i) = \mu(C)$.

So, μ is $\mathbb{Z}_+^\nu$-invariant, belongs to $I_{A\circ f}$ and $f^*(\mu) = \nu$.

Proof of the Lemma. — By eq. (2), we have that,

$$\forall \nu \in I(Y), \qquad \int P(A \circ f, T, f^{-1}(y))d\nu(y) = \int P(0, T, f^{-1}(y))d\nu(y) + \nu(A).$$

Inserting this equality into (2), we have, by assumption (A.4),

$$\sup \{ h_\mu(T) + \mu(A \circ f) \mid f^*(\mu) = \nu \} = h_\nu(S) + \nu(A) + P(0, T, f^{-1}(y))$$

If we take the supremum over $\nu \in I(Y)$ on both sides, we get the result, by the ordinary variational principle, eq. (1).

Proof of Theorem 3 (cfr. [*1*], proof of Theorem 19). — We first remark that

i) $\forall \varepsilon > 0$, $\exists \delta$, such that, $\forall x \in X$, $d(g, g') \leq \delta \rightarrow d(xg, xg') \leq \varepsilon$

ii) $\forall \varepsilon > 0$, $\exists \delta$, such that, $\forall x \in X$, $d(g, g') \geq \varepsilon \rightarrow d(xg, xg') \geq \delta$

i) follows from the uniform continuity of the product $H(x, g) = x.g$ and if *ii*) did not hold, we could have convergent sequences $x_n \rightarrow x$, $g_n \rightarrow g$, $g'_n \rightarrow g'$ with $d(xg, xg') = 0$ and $d(g, g') \geq \varepsilon$; but this would contradict assumption (A.5) *iii*).

Point *i*) and assumptions (A.5) (*v*), imply that, for any $\varepsilon > 0$, there exists a δ, such that, if F is (Λ, δ)-spanning in G, then $x.F$ is (Λ, ε)-spanning in $x.G$, for any $x \in X$. Moreover, by point *ii*), for any $\varepsilon > 0$ there exists a δ, such that, if E is (Λ, ε)-separated in G, then xE is (Λ, δ)-separated (in xG), for any $x \in X$.

Since $f^{-1}(y) = xG$ for some $x \in X$, we get the result by Theorem 1 and the definition of $Z^{(i)}$.

Remark. — In the case where G is a group and where the pressure is G-invariant (e. g. when the representation R acts trivially on G) one can show that f^* is a bijection from $I^{G}_{A \circ f}$ (the set of G-invariant elements of $I_{A \circ f}$) onto I_A. Indeed if μ satisfies,

$$\forall B \in \mathscr{C}(X, \mathbb{R}) \qquad \mu(B) \leq P(A \circ f + B) - P(A \circ f)$$

then

$$\bar{\mu}(B) = \int_G \mu(B \circ g) dg ,$$

with dg the Haar measure on G, satisfies also,

$$\forall B \in \mathscr{C}(X, \mathbb{R}) \qquad \bar{\mu}(B) \leq P(A \circ f + B) - P(A \circ f),$$

due to the G-invariance of P and $A \circ f$.

On the other hand, if two measures $\mu_1, \mu_2 \in I^{G}_{A \circ f}$ have the same images under f^*, they have to coincide on the G-invariant functions, and since they are G-invariant, to coincide everywhere.

3. EXAMPLES

1) Let X, Y be compact metric groups. $\mathbb{Z}^{\nu}_{+}$ acts by continuous endomorphisms on X and f is a continuous epimorphism from X onto Y. Then $G = \operatorname{Ker}(f)$. The two following examples are special cases of this one.

2) We can take $X = K^n$, the n-torus, $Y = K^m$ with $m < n$ and let f be the projection of K^n onto K^m given by the restriction to the first m components. Then G is homeomorphic to K^{n-m} and the topological entropy of an endomorphism on G can be computed (cfr. [*1*]).

3) Our third example is given by classical lattice spin systems. We refer to [*4*] for the terminology.

$X = \{ -1, +1 \}^{\mathbb{L}}$ where $\mathbb{L}$ is a discrete $\mathbb{Z}^{\nu}$-invariant subset of $\mathbb{R}^{\nu}$. For $B \in \mathscr{P}_f(\mathbb{L})$ the character σ_B on X is defined by:

$$\sigma_B(x) = \prod_{i \in B} x_i$$

$\mathscr{B}$ is a $\mathbb{Z}^{\nu}$-invariant family of elements of $\mathscr{P}_f(\mathbb{L})$ and $\mathscr{B}_0$ is a fundamental subfamily of $\mathscr{B}$.

We define a homeomorphism:

$$\gamma = X \to \mathcal{P}(\mathcal{B})$$

by

$$\gamma(x) = \{ B \mid \sigma_B(x) = -1 \}$$

Let $Y = \mathrm{Im}(\gamma)$ (usually denoted Γ) and $f = \gamma$; then $G = \mathrm{Ker}(\gamma)$ (usually denoted $\mathcal{S}$).

For $\alpha, \beta \in \mathcal{P}_f(\mathcal{B})$ let

$$\sigma_\alpha(\beta) = (-1)^{|\alpha \cap \beta|}$$

Given a function $K : \mathcal{B}_0 \to \mathbb{R}$, we define a function $A \in \mathcal{C}(Y, \mathbb{R})$

$$A = -\sum_{B \in \mathcal{B}_0} K(B)\sigma_{\{B\}}$$

Since $\sigma_{\{B\}}(\gamma(X)) = \sigma_B(X)$

$$A \circ \gamma = -\sum_{B \in \mathcal{B}_0} K(B)\sigma_B$$

By the above construction, we see that one can associate with a system $(\mathbb{L}, \mathcal{B}, X, K)$ another system $(\mathbb{L}', \mathcal{B}', X', K')$ with

$$\mathbb{L}' = \mathcal{B}, \mathcal{B}' = \{ \{ B \} \mid B \in \mathcal{B} \}, \qquad X' = \gamma(X) \quad \text{and} \quad K'(\{ B \}) = K(B).$$

The second system is a hard-core system with only one-body interactions (« external field »).

If there is only one equilibrium state for A, then by Theorem 3 and the definition of f^*, all the equilibrium states of $A \circ \gamma$ coincide on the $\mathcal{S}$-invariant functions i. e. for ferromagnetic systems $(K(B) > 0)$, $\Delta^I \subseteq \Delta^+$ (for definitions see [*4*]).

In the present example one can consider equilibrium states which are not necessarily invariant [*5*]. One can show [*6*] that f^* defines a surjection from the set of equilibrium states on X onto the equilibrium states on Y which is a bijection when restricted to $\mathcal{S}$-invariant equilibrium states on X.

ACKNOWLEDGMENT

We thank Prof. J. Slawny for suggesting this work and for very interesting discussions. In particular, Example 3 of Section 3 is due to him.

We thank also the referee and Prof. D. Ruelle for drawing our attention to reference [*7*].

REFERENCES

[*1*] R. Bowen, Entropy for group endomorphisms and homogeneous spaces. *Transact. A. M. S.*, t. **153**, 1971, p. 401-414.

[*2*] D. Ruelle, Statistical mechanics on a compact set with $\mathbb{Z}^\nu$ action satisfying expansiveness and specification. *Transact. A. M. S.*, t. **185**, 1973, p. 237-251.

[3] P. WALTERS, A variational principle for the pressure of continuous transformations. *Amer. Journ. of Math.*, t. **97**, 1975, p. 937-971.
[4] J. SLAWNY, Ferromagnetic spin systems at low temperatures. *Commun. Math. Phys.*, t. **46**, 1976, p. 75-97.
[5] D. RUELLE, Séminaire Bourbaki Année 75/76, n° 480. *Lecture notes in Mathematics*, n° 567 Springer.
[6] J. SLAWNY, *Lectures given at the University of Louvain*, Année 76/77.
[7] F. LEDRAPPIER, P. WALTERS, *A relativised variational principle for continuous transformations*. Preprint.

(Texte révisé reçu le 30 janvier 1978)