Journal of Statistical Physics, Vol. 48, Nos. 5/6, 1987

The Role of Stable Manifolds and Information in the Prigogine Theory of Irreversibility

S. Martinez[1,2] **and E. Tirapegui**[1,2]

Received April 2, 1987

The image of Dirac measures τ_x by the operator Λ of the construction of Prigogine and collaborators is shown to be concentrated in the stable manifold $X^{st}(x)$ and its density function ρ is studied for Bernoulli shifts. The value $v_\infty = \exp[-h_\mu(T)]$, where $h_\mu(T)$ is the Kolmogorov entropy, appears as a critical point for the behavior of ρ. It is also proved that no loss of information is involved by passing from the dynamical system to the Markov process when $v_\infty > 1/2$. The discussion is based on the introduction of an invariant for Markov systems that generalizes the usual Kolmogorov entropy for dynamical systems.

KEY WORDS: K-systems; information; Kolmogorov entropy; stable manifolds; coarse grainings.

1. INTRODUCTION

We study two aspects related to the construction of a Markov process "equivalent" to a K-system proposed recently by Prigogine and collaborators.[1-3] It is a pleasure to dedicate this work to Prof. I. Prigogine, whose influence has been so profound in the attempts to understand the irreversibility of nature and whose ideas are at the origin of the problems treated in this paper.

We recall briefly Prigogine's construction of the Markov process and note that the measurable spaces we consider are Lebesgue. Let $(\Omega, \mathscr{B}, \mu, T)$ be a K-system with generating σ-algebra a_0, $a_n = T^n a_0$, $n \in Z$, $a_n \subset a_{n+1}$, $a_{-\infty} = \{\Omega, \phi\}$, $a_\infty = \mathscr{B}$, μ the invariant normalized measure $\mu(\Omega) = 1$, and

---

[1] Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Casilla 170-3, Santiago, Chile.

[2] Faculté des Sciences, Université Libre de Bruxelles, Belgium.

T the invertible measurable point transformation. Then we construct an invertible operator Λ by

$$\Lambda = \sum_{n \in Z} \lambda_n E_n + R_{-\infty} \tag{1}$$

(where $E_n = R_n - R_{n-1}$, $R_n = E^{a_n}$ is the conditional expectation[4] with respect to a_n taken with μ; then $R_\infty = id$ and $R_{-\infty}$ is a projector over the constants), and $\{\lambda_n \in [0, 1], n \in Z\}$ is a nonincreasing sequence such that $\lambda_n \to 1$, $n \to -\infty$, $\lambda_n \to 0$, $n \to \infty$, and $\{v_n = \lambda_{n+1} \lambda_n^{-1}, n \in Z\}$ is a decreasing sequence. The operator Λ is bounded, Hermitian, and invertible in a dense set of $L^2(\mu)$, and allows the construction of a bounded, doubly stochastic operator $W = \Lambda^{-1} U \Lambda$ nonunitarily equivalent to the operator $Uf = f \circ T$ and given by

$$W = \left(\sum_{n \in Z} \bar{v}_n R_n + v_\infty \right) U \tag{2}$$

where $\bar{v}_n = v_n - v_{n+1} \geqslant 0$, $v_\infty = \lim_{n \to \infty} v_n$, and $0 \leqslant v_\infty < 1$. The Markov system $(\Omega, \mathscr{B}, \mu, Q_W)$ "equivalent" to the K-system is then generated by the transition kernel $Q_W: \Omega \times \mathscr{B} \to [0, 1]$, $Q_W(x, B) = (W 1_B)(x)$, $B \in \mathscr{B}$, $x \in \Omega$, and μ is the invariant measure, since $\int \mu(dx)\, Q_W(x, B) = \mu(B)$. We have proved in Ref. 4 that for K-systems with finite Kolmogorov entropy the measure $Q_W(x, \cdot)$ is strictly concentrated in the stable manifold $X^{\mathrm{st}}(Tx)$ of the transformed point of x by T and that it has a point mass $Q_W(x, \{Tx\}) \geqslant v_\infty$ if $v_\infty > 0$, with the equality holding for Bernoulli shifts.

The first aspect we discuss here refers to the interpretation of the operator Λ. We have argued in Ref. 5 that the elementary objects whose evolution in time should be considered are probability densities $\hat{\delta}_x$ with support in the stable manifolds of points $x \in \Omega$, since these points share the same behavior in the future (see also Ref. 6 in this context). We must look then for an operator Λ transforming the elementary probability density evolving with U^* which is δ_x, a Dirac δ-function centered in x, in $\hat{\delta}_x = \Lambda \delta_x$. The simplest possible Λ satisfying this requirement is given by (1) and then the observable evolution will be given by $W^* = \Lambda U^* \Lambda^{-1}$. Here U^* is the adjoint of U in $L^2(\mu)$, which acts as $(U^* f)(x) = f(T^{-1} x)$ on bounded measurable functions and W^*, the adjoint of W in $L^2(\mu)$, gives the observable evolution of probability densities in the Markov system $(\Omega, \mathscr{B}, \mu, Q_W)$. In relation to this interpretation of Λ we prove in Section 2 that for K-shifts the measure $Q_\Lambda(x, B) = (\Lambda 1_B)(x)$, $B \in \mathscr{B}$, induced by Λ (whose density is $\hat{\delta}_x$) is concentrated in the stable manifold $X^{\mathrm{st}}(x)$ of x, i.e., $Q_\Lambda(x, X^{\mathrm{st}}(x)) = 1$, and when $\lambda_n = 1$, $n \leqslant q$, it is concentrated in

$$B(-q, \infty)(x) = \{ y \in \Omega : y_j = x_j, j \geqslant -q \} \subset X^{\mathrm{st}}(x)$$

For Bernoulli shifts in this last case we have that $Q_\Lambda(x, \cdot)$ is absolutely continuous in $B(-q, \infty)(x)$ with respect to the measure induced by μ in this set with density ρ. If for $n \geqslant q$ one has $\lambda_n = c^{n-q}$, $c < 1$, then $v_\infty = c$ and ρ is bounded for $c < \exp[-h_\mu(T)]$ and unbounded for $c > \exp[-h_\mu(T)]$, where $h_\mu(T)$ is the Kolmogorov entropy of the shift. When $\{\lambda_n, n \geqslant q\}$ decreases as $e^{-\phi(n)}$, with $\phi(n)$ growing faster than n, ρ is bounded.

The second aspect we treat (Section 3) is an attempt to explain in what sense there is no loss of information when passing from the K-system to the Markov system Q_W. We define an invariant $\bar{h}_\mu$ of Markov systems, which can be interpreted as the maximum loss of information in one step of the Markov system when the process is regarded through all possible future σ-algebras. The invariant $\bar{h}_\mu$ is a generalization to Markov systems of the usual Kolmogorov entropy of dynamical systems. We show that for the process Q_W one has $\bar{h}_\mu(Q_W) = h_\mu(T)$ if $v_\infty > 1/2$ and that this equality is a consequence of the fact that if one regards the system considering only events in any of the σ-algebras a_n, then two fibers ξ and ξ' in a_n have the same future with the Markov evolution if and only if their future with the dynamical system is the same. These results show, then, in what sense there is no loss of information when the evolution is generated by the Markov process Q_W instead of T. Moreover, if $\bar{h}_\mu$ is evaluated for a class of coarse grainings obtained through projections generated by future σ-algebras, it turns out to be infinite.

2. PROPERTIES OF THE MEASURE INDUCED IN THE STABLE MANIFOLDS

We study here the density $\hat{\delta}_x = \Lambda \delta_x$ image by Λ of the δ-function δ_x centered at x, which corresponds to the measure τ_x concentrated in $\{x\}$. The function $\hat{\delta}_x$ will be the density of the measure $Q_\Lambda(x, \cdot)$, where $Q_\Lambda: \Omega \times \mathscr{B} \to [0, 1]$ is the transition kernel generated by the Hermitian, doubly stochastic operator Λ given by (1). It is this measure $Q_\Lambda(x, \cdot)$ that is the object of interest to us. Let the dynamical system $(\Omega, \mathscr{B}, \mu, T)$ now be a Bernoulli shift; then $\Omega = S^Z$, where $S = \{1,..., d\}$ is the alphabet of the shift and μ is the product measure induced in S^Z by the probability vector $(\Pi_1 \cdots \Pi_d)$ in S. The transformation T is the shift $(Tx)_n = x_{n+1}$ and the generating σ-algebra a_0 is $\bigvee_{j \leqslant 0} T^j \alpha$, where α is the partition $\alpha = \{X_i: x \in X_i \Leftrightarrow x_0 = i\}$. The stable manifold $X^{\mathrm{st}}(x)$ of $x \in \Omega$ is the disjoint union

$$X^{\mathrm{st}}(x) = \left[\bigcup_{k \in Z} \hat{X}_k^{\mathrm{st}}(x)\right] \bigcup \{x\}$$

where

$$\hat{X}_k^{\mathrm{st}}(x) = \{z \in \Omega : z_j = x_j,\ j \geqslant k,\ z_{k-1} \neq x_{k-1}\}$$

Using the method in Ref. 4, one finds ($\bar{\lambda}_n \equiv \lambda_n - \lambda_{n+1}$):

$$Q_k \equiv Q_\Lambda(x, \hat{X}_k^{\mathrm{st}}(x))$$
$$= [\Pi_{k-1}(x)^{-1} - 1] \sum_{u=0}^{\infty} \Pi_{k-1}(x)\, \Pi_k(x) \cdots \Pi_{k+u-1}(x)\, \bar{\lambda}_{-(k+u)}$$

where we use the notation $\Pi_k(x) \equiv \Pi_{x_k}$. One easily checks $\sum_{k \in Z} Q_k = 1$ and then $Q_\Lambda(x, X^{\mathrm{st}}(x)) = 1$. The disjoint sets $\hat{X}_k^{\mathrm{st}}(x)$ contain points that are more and more different from x when k increases, and consequently with any reasonable distance defined in the space Ω their average distance to x will increase with $k \in Z$. Then, in agreement with the physical interpretation we have proposed for Λ in Ref. 5, we can make a choice for the sequence (λ_n) such that the weights Q_k are zero for k bigger than some number in Z. If we choose $\lambda_n = 1$, $n \leqslant q$, $\lambda_n < 1$, $n > q$, one easily checks that

$$Q_{-q+n} = 0, \quad n \geqslant 1; \qquad Q_{-q-n} \neq 0, \quad n \geqslant 0 \tag{3}$$

i.e., the measure $Q_\Lambda(x, \cdot)$ is concentrated in

$$B(-q, \infty)(x) = \{z \in \Omega;\ z_j = x_j,\ j \geqslant -q\}$$

and this set is the disjoint union

$$B(-q, \infty)(x) = \left[\bigcup_{n \geqslant 0} \hat{X}_{-q-n}^{\mathrm{st}}(x)\right] \bigcup \{x\} \tag{4}$$

On the other hand, it follows from our discussion in Ref. 4 that what is finally important is the behavior of λ_n for big n, since this determines the value of $v_\infty < 1$. In the case $v_\infty > 0$ one has that λ_n behaves asymptotically, when $n \to \infty$, as c^n with $c < 1$, and $v_\infty = c$. In the other case $v_\infty = 0$ and this means that, for big n, λ_n behaves as $e^{-\phi(n)}$, where $\phi(n)$ grows with n faster that n. We can then model in a simple way the first situation, taking

$$\lambda_n = 1, \quad n \leqslant q; \qquad \lambda_n = c^{n-q}, \quad n \geqslant q \tag{5}$$

We study this case in what follows. One has $Q_{-q+n} = 0$, $n \geqslant 1$, and

$$Q_{-q-n} = [\Pi_{-q-n-1}(x)^{-1} - 1](1-c)$$
$$\times \sum_{r=0}^{n} c^{n-r} \Pi_{-q-n-1}(x)\, \Pi_{-q-n}(x) \cdots \Pi_{-q-n+r-1}(x) \tag{6}$$

for $n \geqslant 0$. The measure $Q_A(x, \cdot)$ is then strictly concentrated in $B(-q, \infty)(x)$. Let $\bar{\mathscr{B}}$ be the σ-algebra induced by $\mathscr{B}$ in $B(-q, \infty)(x)$; it is generated by the sets

$$B(-q-r, \infty)(x) = \{z: z_{-q-j} = i_j,\ 1 \leqslant j \leqslant r;\ z_j = x_j,\ j \geqslant -q\} \tag{7}$$

where $i_j \in S$, the alphabet of the shift. The invariant Bernoulli measure μ induces a measure $\mu_{-q,x}$ in the space $(B(-q, \infty)(x), \bar{\mathscr{B}})$ defined by

$$\mu_{-q,x}(B(-q-r, \infty)(x)) = \Pi_{i_1} \Pi_{i_2} \cdots \Pi_{i_r} \tag{8}$$

We want to see now how $Q_A(x, \cdot)$ is distributed in $(B(-q, \infty)(x), \bar{\mathscr{B}})$ and for this we define the sets $(n \geqslant 0)$

$$\begin{aligned} &\hat{B}(-q-n-r, \infty)(i_1 \neq x_{-q-n-1}, i_2, \ldots, i_r) \\ &\quad = \{z: z_{-q-n-j} = i_j,\ 1 \leqslant j \leqslant r;\ z_k = x_k,\ k \geqslant -q-n\} \end{aligned} \tag{9}$$

These sets are contained in $\hat{X}^{\mathrm{st}}_{-q-n}(x)$ and they generate the σ-algebra $\bar{\mathscr{B}}_{-q-n}$ induced by $\bar{\mathscr{B}}$ in $\hat{X}^{\mathrm{st}}_{-q-n}(x)$. One has

$$\begin{aligned} &\mu_{-q,x}(\hat{B}(-q-n-r, \infty)(x)) \\ &\quad = \Pi_{i_1} \cdots \Pi_{i_r} \Pi_{-q-n}(x)\, \Pi_{-q-n+1}(x) \cdots \Pi_{-q-1}(x) \end{aligned} \tag{10}$$

$$\begin{aligned} &Q_A(x, \hat{B}(-q-n-r, \infty)) \\ &\quad = \Pi_{i_1} \cdots \Pi_{i_r} \Pi_{-q-n-1}(x)^{-1} \\ &\qquad \times \sum_{u=0}^{n} \Pi_{-q-n-1}(x)\, \Pi_{-q-n}(x) \cdots \Pi_{-q-n-1+u}(x)\, c^{n-u}(1-c) \end{aligned} \tag{11}$$

Formulas (10) and (11) show that $Q_A(x, \cdot)$ is absolutely continuous with respect to $\mu_{-q,x}$ in the measure space $(\hat{X}^{\mathrm{st}}_{-q-n}(x), \bar{\mathscr{B}}_{-q-n})$ and its density there takes a constant value $\bar{Q}_{-q-n}$ given by

$$\bar{Q}_{-q-n} = (1-c)\left(1 + \sum_{j=1}^{n} \frac{c^j}{\Pi_{-q-1}(x)\, \Pi_{-q-2}(x) \cdots \Pi_{-q-j}(x)}\right) \tag{12}$$

for $n \geqslant 1$, $\bar{Q}_{-q} = 1 - c$. An equivalent statement to this last one for the transition probability measure $Q_W(x, \cdot)$ of the Markov process was given by Proposition 1 of Ref. 4. We conclude then that the measure $Q_A(x, \cdot)$ is absolutely continuous with respect to $\mu_{-q,x}$ in the space $(B(-q, \infty)(x), \bar{\mathscr{B}})$ and its density $\rho \in L^1(\mu_{-q,x})$ there is

$$\rho = \sum_{n \geqslant 0} \bar{Q}_{-q-n}\, 1_{\hat{X}^{\mathrm{st}}_{-q-n}(x)} \tag{13}$$

$$\int_{B(-q,\infty)(x)} \rho \, d\mu_{-q,x} = \sum_{n \geqslant 0} Q_{-q-n} = 1 \tag{14}$$

From (13) we see that ρ is a step function in $B(-q, \infty)(x)$ taking an infinite number of values there. The amplitudes of the jumps of ρ are

$$\Delta \bar{Q}_n = \bar{Q}_{-q-n} - \bar{Q}_{-q-n+1} = \frac{(1-c)\, c^n}{\Pi_{-q-1}(x) \cdots \Pi_{-q-n}(x)} \tag{15}$$

In order to study in more detail the properties of ρ, we put

$$a_j(x) = \Pi_{-q-1}(x)\, \Pi_{-q-2}(x) \cdots \Pi_{-q-j}(x), \qquad j \geqslant 1$$

Then

$$\bar{Q}_{-q-n} = (1-c)\, v_n(x), \qquad v_n(x) = \sum_{j=0}^{n} \frac{c^j}{a_j(x)}, \qquad n \geqslant 0 \tag{16}$$

From (13) we see that the function ρ will be bounded if $v_n(x)$ is uniformly bounded for all n and unbounded if this is not the case. From the Shannon–McMillan–Breiman theorem[7] one knows that one can extract from the space Ω a set N of μ measure zero such that $\forall x \in \Omega \backslash N$ one has for any r

$$\lim_{n \to \infty} -\frac{1}{n} \log \Pi_r(x)\, \Pi_{r-1}(x) \cdots \Pi_{r-n+1}(x) = h_\mu(T) \tag{17}$$

We restrict ourselves to $\Omega \backslash N$. Then $\forall \varepsilon > 0$, $\exists n_x(\varepsilon)$ such that

$$e^{-n(h_\mu(T)+\varepsilon)} \leqslant a_n(x) \leqslant e^{-n(h_\mu(T)-\varepsilon)}, \qquad \forall n \geqslant n_x(\varepsilon) \tag{18}$$

Putting $t_n(x) = a_n(x) \exp[nh_\mu(T)]$, one also has

$$\exp\left[-\frac{1}{n} \log t_n(x)\right] \to 1 \qquad \text{if} \quad n \to \infty \tag{19}$$

Let $c < \exp[-h_\mu(T)] \equiv s$; then $\exists \varepsilon > 0$ such that $ce^\varepsilon < s$, and due to (18), $\exists n_x(\varepsilon)$ such that $\forall j \geqslant n_x(\varepsilon)$ one has $a_j(x) > s^j \exp(-j\varepsilon)$. Then

$$\sum_{j \geqslant n_x(\varepsilon)} \frac{c^j}{a_j(x)} < \sum_{j > n_x(\varepsilon)} \left(\frac{ce^\varepsilon}{s}\right)^j = \left(\frac{ce^\varepsilon}{s}\right)^{n_x(\varepsilon)} \frac{1}{1 - ce^\varepsilon/s} = K_0 \tag{20}$$

Let $\underline{\Pi} = \min \Pi_i$, $\bar{\Pi} = \sup \Pi_i$, $0 < \underline{\Pi} \leqslant \bar{\Pi} < 1$. Then

$$\sum_{j=0}^{n_x(\varepsilon)-1} \frac{c^j}{a_j(x)} \leqslant \sum_{j=0}^{n_x(\varepsilon)-1} \left(\frac{c}{\underline{\Pi}}\right)^j = K_1 \tag{21}$$

From these last two inequalities and the expression (16) for $v_n(x)$ we conclude $v_n(x) < K_0 + K_1$ for all n and consequently ρ is a bounded function when $c < s$. An immediate consequence of this is that $\rho \in L^p(\mu_{-q,x})$, $\forall p \geqslant 1$.

If $c > s$, one can write $c = se^{\eta}$, $\eta > 0$, and if $0 < \varepsilon < \eta$, we can choose $n_x(\varepsilon)$ such that

$$v_n(x) > \frac{c^n}{a_n(x)} > e^{n(\eta - \varepsilon)}, \qquad \forall n \geqslant n_x(\varepsilon) \tag{22}$$

which shows that ρ is unbounded. We consider now

$$\int \rho^p \, d\mu_{-q,x} = (1-c)^p \sum_{n \geqslant 0} v_n(x)^p \, [1 - \Pi_{-q-n}(x)] \, a_n(x) \tag{23}$$

Since

$$\mu_{-q,x}(\hat{X}^{\mathrm{st}}_{-q-n}(x)) = [1 - \Pi_{-q-n-1}(x)] \, a_n(x)$$

One has

$$1 - \bar{\Pi} \leqslant 1 - \Pi_{-q-n-1}(x) \leqslant 1 - \underline{\Pi}$$

and consequently the integral in (23) will converge if

$$I_p = \sum_{n \geqslant 0} u_n(x) \, s^n < \infty, \qquad u_n(x) = v_n(x)^p \, t_n(x) \tag{24}$$

where $t_n(x)$ is defined after (18). The radius of convergence γ of the power series in (24) is

$$\gamma = \lim_{n \to \infty} u_n(x)^{-1/n} = \lim_{n \to \infty} \exp\left[-\frac{p}{n} \log v_n(x) \right] \tag{25}$$

due to (19). One has

$$v_n(x) > \frac{c^n}{a_n(x)} = \left\{ \frac{c}{s} \exp\left[-\frac{1}{n} \log t_n(x) \right] \right\}^n \tag{26}$$

and then, using (19), we obtain

$$\gamma \leqslant \exp[-p \log(cs^{-1})] \tag{27}$$

The series in (24) will diverge if $s > \gamma$. Using (27), this will be satisfied if

$$p > p_0 \equiv -\frac{\log s}{\log cs^{-1}} \tag{28}$$

where $p_0 > 1$ since $c > s$. We see then that given $c > s$, there exists a number $p_0 > 1$ such that $\rho \notin L^p(\mu_{-q,x})$ for $p > p_0$. We have then proved the following result.

Theorem 1. Let $(\Omega, \mathscr{B}, \mu, T)$ be a Bernoulli shift and $\Lambda = \sum_{n\in Z} \lambda_n E_n + R_{-\infty}$, $\lambda_n = 1$, $n \leqslant q$, and $\lambda_n = c^{n-q}$, $n \geqslant q$, with $c < 1$; then one has:

(a) The transition kernel $Q_\Lambda(x, \cdot)$ is strictly concentrated in the set $B(-q, \infty)(x) \subset X^{st}(x)$, the stable manifold of x, where

$$B(-q, \infty)(x) = \{x' \in \Omega : x'_j = x_j, j \geqslant -q\}$$

(b) The measure $Q_\Lambda(x, \cdot)$ is absolutely continuous with respect to the measure $\mu_{-q,x}$ induced in $B(-q, \infty)(x)$ by μ and its density ρ has the following properties: (i) If $c < \exp[-h_\mu(T)]$, where $h_\mu(T)$ is the Kolmogorov entropy of the shift, then ρ is bounded and $\rho \in L^p(\mu_{-q,x})$ for all $p \geqslant 1$; (ii) if $c > \exp[-h_\mu(T)]$, then ρ is unbounded and there exists a number $p_0 > 1$ such that $\rho \notin L^p(\mu_{-q,x})$ for $p > p_0$.

Let us remark that this result gives no information when $c = \exp[-h_\mu(T)]$. However, it is easy to see that if $\Pi_1 = \cdots = \Pi_d = d^{-1}$, then ρ is unbounded but belongs to $L^p(\mu_{-q,x})$ for $p \geqslant 1$. The above theorem refers to the case $v_\infty > 0$; when $v_\infty = 0$ this means that λ_n behaves for big n as $e^{-\phi(n)}$ with $\phi(n)$ growing faster than n and then it is simple to see that ρ will be bounded. One can also check, using again the method in Ref. 4 to describe $Q_\Lambda(x, \cdot)$, that part (a) of the theorem remains valid for K-shifts.

3. CONSERVATION OF INFORMATION AND A RELATED INVARIANT FOR MARKOV SYSTEMS

We introduce in this section an invariant for Markov systems that can be interpreted as the maximal possible loss of information in one step of the process. Let $(\Omega, \mathscr{B})$ be a measurable space and Q a transition kernel $Q: \Omega \times \mathscr{B} \to [0, 1]$, $(x, B) \to Q(x, B) \in [0, 1]$; then Q acts on the bounded measurable functions $(Qf)(x) = \int Q(x, dx') f(x')$ and on the measures $(vQ)(B) = \int v(dx') Q(x', B)$ and one has $Q(x, B) = (Q1_B)(x)$. If μ is an invariant measure, i.e., $\mu Q = \mu$, then we call $(\Omega, \mathscr{B}, \mu, Q)$ a Markov system.

Let $a \subset \mathscr{B}$ be a σ-algebra and let us define a new σ-algebra $Q^{(-1)}a$ as the smallest σ-algebra that makes all the functions $\{Q1_A, A \in a\}$ measurable. If $Q^{(-1)}a \subset a$, we say that a is a Q-invariant σ-algebra and in this case we can interpret a as a future σ-algebra in the sense that two points in the same fiber of a (a fiber of a is an equivalence class generated by the equivalence relation $x \cong x'$ iff x and x' are not separated by any set of a) will have the same future when only events in a can be observed. In order to see this, let x and x' belong to the same fiber $\xi(x) = \xi(x')$ of a [the notation $\xi(x)$ stands for the fiber of a containing x]; then we cannot distinguish them at time zero and, moreover, $Q^n(x, A) = Q^n(x', A)$, $\forall A \in a$, $\forall n \geq 1$, where $Q^n(x, A) = Q^n 1_A(x)$ is the transition probability at time n. We note that $Q^{(-1)}a \subset a$ means in fact that the transition kernel Q can be considered as a transition kernel in the space $\hat{a}$ of fibers of a, i.e., $Q: \hat{a} \times a \to [0, 1]$, $Q(\xi(x), A) = Q(x, A)$. Let us consider now two points x and x' in the same fiber of $Q^{(-1)}a$; then $Q^n(x, A) = Q^n(x', A)$, $\forall A \in a$, $\forall n \geq 1$, and they cannot be distinguished by events in a for times $n \geq 1$, although they can be distinguished at time zero, since, if $Q^{(-1)}a \subset a$, they can belong to different fibers of a. We can then interpret the conditional entropy (with respect to the invariant measure μ) $H_\mu(a \mid Q^{(-1)}a)$ as the loss of information in one time step when a is the σ-algebra of observable events. We can now define

$$\bar{h}_\mu(Q) = \sup_{Q^{(-1)}a \subset a} H_\mu(a \mid Q^{(-1)}a) \tag{29}$$

and this quantity will represent the maximum loss of information in one step when the system is regarded with all possible future σ-algebras. The quantity $\bar{h}_\mu(Q)$ was introduced and studied in Ref. 8; it is an invariant by isomorphisms of Markov systems and it is a generalization of the Kolmogorov entropy for dynamical systems. Indeed, if $(\Omega, \mathscr{B}, \mu, T)$ is a dynamical system, then it is a Markov system with the transition kernel

$$T(x, B) = (U 1_B)(x) = 1_{T^{-1}B}(x)$$

where $U = f \circ T$ will be denoted simply by T when there is no risk of confusion, and one has $T^{(-1)}a = T^{-1}a$ for any σ-algebra $a \subset \mathscr{B}$; then (29) is the well-known formula of Rokhlin[9] for the Kolmogorov entropy $h_\mu(T)$.

Let us consider now the dynamical system $(\Omega, \mathscr{B}, \mu, T)$ and some Markov systems related to it, for which we shall study the invariant (29). Let a be a T-invariant σ-algebra, i.e., $T^{-1}a \subset a \subset \mathscr{B}$; then the operator $\bar{W} = E^a U$, where E^a is the conditional expectation with respect to a taken with μ, is doubly stochastic and generates a Markov system $(\hat{a}, a, \mu, Q_{\bar{W}})$, where $\hat{a}$ is the space of fibers of a, a denotes by an abuse of notation the σ-algebra induced by a in $\hat{a}$, and the transition kernel $Q_W: \hat{a} \times a \to [0, 1]$ is given by $Q_{\bar{W}}(\xi(x), A) = (\bar{W} 1_A)(x)$, $\xi(x)$ the fiber of a containing x. We call this Markov system a coarse graining. Let $\tilde{a} \subset a$ be a T-invariant σ-algebra, $T^{-1}\tilde{a} \subset \tilde{a}$; then

$$E^a U 1_A = 1_{T^{-1}A}, \qquad A \in \tilde{a}$$

and $(E^a U)^{(-1)}\tilde{a} = T^{-1}\tilde{a}$. We have then

$$H_\mu(\tilde{a} \mid (E^a U)^{(-1)}\tilde{a}) = H_\mu(\tilde{a} \mid T^{-1}\tilde{a})$$

and consequently

$$\bar{h}_\mu(Q_{\bar{W}}) = \operatorname*{Sup}_{\bar{W}^{(-1)}\tilde{a} \subset \tilde{a} \subset a} H_\mu(\tilde{a} \mid T^{-1}\tilde{a}) = h_\mu(T_{\hat{a}})$$

where $h_\mu(T_{\hat{a}})$ is the Kolmogorov entropy of the dynamical system $(\hat{a}, a, \mu, T_{\hat{a}})$ induced by the original one in the space of fibers $(\hat{a}, a)$. Since $h_\mu(T_{\hat{a}}) \leqslant h_\mu(T)$, one has $\bar{h}_\mu(Q_{\bar{W}}) \leqslant h_\mu(T)$. If a is of full entropy, i.e., $h_\mu(T) = H_\mu(a \mid T^{-1}a)$, then $\bar{h}_\mu(Q_{\bar{W}}) = h_\mu(T_{\hat{a}}) = h_\mu(T)$.

We consider now the Markov system $(\Omega, \mathscr{B}, \mu, Q'_{\bar{W}})$ induced by $\bar{W} = E^a U$ on the original space $(\Omega, \mathscr{B})$. If the T-invariant σ-algebra $\tilde{a}$ is finer than a, $a \subset \tilde{a}$, then, since $E^a U 1_A$ is a-measurable, one has $(E^a U)^{(-1)}\tilde{a} \subset a$ and consequently

$$H_\mu(\tilde{a} \mid (E^a U)^{(-1)}\tilde{a}) \geqslant H_\mu(\tilde{a} \mid a)$$

If $T^{-1}a$ is strictly contained in a, we consider the increasing family of σ-algebras $(a_n = T^n a,\ n \geqslant 0)$. Then

$$H_\mu(a_n \mid (E^a U)^{(-1)}a_n) \geqslant H_\mu(a_n \mid a) = nH_\mu(a_1 \mid a_0), \qquad n \geqslant 1$$

and consequently $\bar{h}_\mu(Q'_{\bar{W}}) = \infty$. We have proved the following result:

Proposition 1. Let $(\Omega, \mathscr{B}, \mu, T)$ be a dynamical system and a a T-invariant σ-algebra. Then:

(a) The Markov system $(\hat{a}, a, \mu, Q_{\bar{W}})$ induced in the space $\hat{a}$ of the fibers of a by the doubly stochastic operator $\bar{W} = E^a U$ is such that the invariant $\bar{h}_\mu(Q_{\bar{W}}) = h_\mu(T_{\hat{a}}) \leqslant h_\mu(T)$, where $h_\mu(T_{\hat{a}})$ is the Kolmogorov entropy of the dynamical system $(\hat{a}, a, \mu, T_{\hat{a}})$ induced by T in $(\hat{a}, a)$ and $h_\mu(T)$ is the Kolmogorov entropy of T. If a is full entropy, $\bar{h}_\mu(Q_{\bar{W}}) = h_\mu(T)$.

(b) The Markov system $(\Omega, \mathscr{B}, \mu, Q'_{\bar{W}})$ induced by $\bar{W}$ in the original space $(\Omega, \mathscr{B})$ is such that $\bar{h}_\mu(Q'_{\bar{W}}) = \infty$ if $T^{-1}a$ is strictly contained in a.

If our dynamical system is now a K-system $(\Omega, \mathscr{B}, \mu, T)$ with generating σ-algebra a_0, $a_n = T^n a_0$, $n \in Z$, then we can associate to it the Markov system $(\Omega, \mathscr{B}, \mu, Q_W)$ through the Λ operator of Section 1. The transition probability Q_W is constructed with the operator W given by (2), and if $A \in a_m$, one has

$$W 1_A = \sum_{n \leqslant m-2} \bar{v}_n E^{a_n} 1_{T^{-1}A} + v_{m-1} 1_{T^{-1}A}$$

which is a_{m-1} measurable and consequently one has $Q_W^{(-1)} a_m \subset a_{m-1}$, which implies

$$H_\mu(a_m \mid Q_W^{(-1)} a_m) \geqslant H_\mu(a_m \mid a_{m-1}) = h_\mu(T)$$

and then $\bar{h}_\mu(Q_W)$ defined by (29) satisfies $\bar{h}_\mu(Q_W) \geqslant h_\mu(T)$. We shall see that the equality holds in some special cases.

Putting

$$Q''_W = \sum_{n \in Z} \bar{v}_n (1 - v_\infty)^{-1} E^{a_n} U$$

we can write $W = v_\infty U + (1 - v_\infty) Q''_W$. We consider now a more general situation. Let Q'' be a transition kernel in $(\Omega, \mathscr{B})$ and $U_S f = f \circ S$, where S is a point transformation in Ω. Then, for $0 \leqslant \gamma \leqslant 1$, the double stochastic operator $\bar{Q} = \gamma U_S + (1 - \gamma) Q''$ generates a Markov system $(\Omega, \mathscr{B}, \mu, \bar{Q})$ if the measure is both S-invariant and Q''-invariant. In what follows we assume $\gamma > 1/2$. Then $S^{-1}B = \{x: (\bar{Q} 1_B)(x) > 1/2\}$, $B \in \mathscr{B}$. Let $a \subset \mathscr{B}$ be a σ-algebra; then, since $\bar{Q} 1_A$ is $\bar{Q}^{(-1)}a$-measurable for any $A \in a$, one has that $S^{-1} = \{x: (\bar{Q} 1_A)(x) > 1/2\}$ is $\bar{Q}^{(-1)}a$-measurable and consequently $S^{-1} a \subset \bar{Q}^{(-1)} a$. This means that any σ-algebra a that is $\bar{Q}$-invariant, i.e., $\bar{Q}^{(-1)} a \subset a$, is also S-invariant, since $S^{-1} a \subset \bar{Q}^{(-1)} a \subset a$, and then one has $H_\mu(a \mid \bar{Q}^{(-1)} a) \leqslant H_\mu(a \mid S^{-1} a)$. This implies $\bar{h}_\mu(\bar{Q}) \leqslant h_\mu(S)$. Since our operator W is of the form of $\bar{Q}$ for $\gamma = v_\infty$, we conclude that for $v_\infty > 1/2$ one has $\bar{h}_\mu(Q_W) \leqslant h_\mu(T)$ and consequently we have in this case $\bar{h}_\mu(Q_W) = h_\mu(T)$. We have proved the following.

Theorem 3. The Markov system $(\Omega, \mathscr{B}, \mu, Q_W)$ associated to the K-system $(\Omega, \mathscr{B}, \mu, T)$ through the doubly stochastic operator $W = (\sum_{n \in Z} \bar{v}_n E^{a_n} + v_\infty) U$, $Q_W(x, B) = (W 1_B)(x)$, is such that the invariant $\bar{h}_\mu(Q_W) = h_\mu(T)$, the Kolmogorov entropy of T, when $v_\infty > 1/2$. In all other cases one has $\bar{h}_\mu(Q_W) \geqslant h_\mu(T)$.

It is important to remark that in the case $v_\infty > 1/2$ the proof of the above theorem implies that $Q_W^{(-1)} a_n = T^{-1} a_n = a_{n-1}$, $\forall n \in Z$, and this means that if we consider a_n as the σ-algebra of observable events, then any two fibers ξ and ξ' of $\hat{a}_n$ generate the same future Markov process if and only if their future with the evolution of the dynamical system is the same.

4. CONCLUSIONS

We discuss first the process Q_W of Section 2, where $\lambda_n = 1$, $n \leqslant q$, $\lambda_n = c^{n-q}$, $n \geqslant q$, $v_\infty = c < 1$. We have seen that in this case the value $c = \exp[-h_\mu(T)] \equiv s$ was a critical one in the sense that the density ρ of the measure $Q_A(x, \cdot)$ in $B(-q, \infty)(x) \subset X^{\mathrm{st}}(x)$ was bounded for $c < s$ and unbounded for $c > s$. In fact, this choice of Λ indicates that we should compare the resulting process with the one obtained by a coarse-graining $R_{q-1} U$ in the sense of Section 3, which amounts to projecting the density functions with R_q, since $R_q U R_q = R_{q-1} U$, and this means that we are taking a_q as a σ-algebra of the future, since projecting with $R_q = E^{a_q}$ says that we are identifying points in the same fiber of a_q as having the same observable future (this class of coarse-grainings was discussed in Ref. 10 in relation to the definition of a nonequilibrium entropy). The difference with our Λ is that the measure $\hat{Q}(x, \cdot)$ defined by $\hat{Q}(x, A) = (R_q 1_A)(x)$ is concentrated uniformly in $(B(-q, \infty)(x), \mathscr{B})$ with density one, while $\Lambda\delta_x$ has density ρ and we have seen then that the only case that is similar to the coarse-graining, which has a clear physical interpretation, is when ρ is bounded, i.e., $c < s$. Moreover, the interpretation of a_q as a future σ-algebra tells us that if x' and x'' are in the same fiber of a_q then $\Lambda\delta_{x'}$ and $\Lambda\delta_{x''}$ should not differ very much and this is true only if $c < s$. What we have gained using Λ instead of R_q is that Λ is an invertible operator and then we can argue that it transforms the evolution of the original dynamical system to an equivalent representation without loss of information, a point discussed in Section 3. Let us recall that the quantity $\nu_\infty = c$ has a simple interpretation in terms of observability of the deterministic trajectory with the Markov process Q_W[11], since the probability of the trajectory $(x, Tx,..., T^t x)$ is ν_∞^t; then, putting $\nu_\infty = \exp(-\tau^{-1})$, we have that the time τ of observation satisfies $\tau < h_\mu(T)^{-1}$ when $c < s$. We remark that for the model of Section 2 the operator W takes the simple form $W = cU + (1-c)\, R_{q-1} U$, showing explicitly the relation to the coarse-graining obtained in the limit $c \to 0$.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the Fondo Nacional de Ciencias, Chile, the DIB of the Universidad de Chile, and the Instituts Internationaux de Physique et Chimie Solvay, Brussels. They also thank Profs. G. Nicolis and I. Prigogine for stimulating discussions and for their hospitality at the University of Brussels.

REFERENCES

1. B. Misra, I. Prigogine, and M. Courbage, *Physica* **98A**:1 (1979).
2. B. Misra and M. Courbage, *Physica* **104A**:359 (1980).
3. S. Goldstein, B. Misra, and M. Courbage, *J. Stat. Phys.* **25**:111 (1981).
4. S. Martinez and E. Tirapegui, *J. Stat. Phys.* **37**:173 (1984).
5. S. Martinez and E. Tirapegui, *Phys. Lett. A* **110**:81 (1985).
6. Y. Elskens and I. Prigogine, *Proc. Natl. Acad. Sci. USA* **63**:5756 (1986).
7. K. Petersen, *Ergodic Theory*, Chapter VI (Cambridge University Press, 1983).
8. S. Martinez, Thèse, Université de Paris VI (1986).
9. V. A. Rokhlin, *Russ. Math. Surv.* **22**(5):1 (1967).
10. S. Goldstein and O. Penrose, J. Stat. Phys. **24**:325 (1981).
11. S. Martinez and E. Tirapegui, in *Irreversible Evolution of Dynamical Systems* (Lecture Notes in Physics, Vol. 179), L. Garrido, ed. (Springer, 1982).