Evolution of ensembles toward an absorbing attractor

A. Nieva
Departamento de Matemáticas, Facultad de Ciencias
Universidad Nacional Autónoma de México
Circuito Exterior de C.U. Coyoacán, 04510 México, D.F., México

Recibido el 27 de febrero de 1992; aceptado el 3 de junio de 1993

Abstract. In Ref. [21] the evolution of ensembles of states for dynamical systems (DS) with states in a probability space is studied. Enlarging some results in Ref. [21] to a very common measure spaces the above study is generalized for DSs with states in those measure spaces. This enlargement uncovers an indecomposability subregion (AA) of the space of states, for which each orbit of states of the system has all its states in AA from some time on. It is obtained that each ensemble of states, in its natural evolution, evolves in ergodic sense (convergence *à la weak* Cesàro) to the microcanonical ensemble (scattered by all AA). A particular but central kind of that evolution is the irreversible one in coarse-grained sense (weak convergence) to the microcanonical ensemble; it is obtained that such evolution holds if and only if the DS is mixing for measure spaces. For this kind of DS predictability [24] is impossible beyond a certain time, however the only forecasting in coarse-grained sense about the system's future which can be made is that of a statistical nature.

Resumen. En la Ref. [21] se analiza la evolución de colectivos de estados para sistemas dinámicos (SD) con estados en un espacio de probabilidad. Extendiendo a espacios de medida algunos resultados en la Ref. [21], en el presente artículo generalizamos ese análisis a colectivos de estados para SD con estados en espacios de medida muy comunes. Nuestra extensión revela una subregión (AA) indescomponible, del espacio de estados, en la cual cada órbita a partir de cierto tiempo tiene todos sus estados en AA. Se obtiene que en su evolución natural cada colectivo de estados tiende hacia el colectivo microcanónico (distribuido en AA) en forma ergódica (convergencia *à la* Cesàro débil). Un tipo particular pero importante de esta evolución es la evolución irreversible en grano grueso hacia el colectivo microcanónico; se obtiene que esta evolución vale si y sólo si el SD es mezclante para espacios de medida. Para esta clase de SD la predicción [24] es imposible más allá de cierto tiempo, sin embargo los únicos pronósticos en grano grueso que se pueden hacer sobre el futuro del sistema son de naturaleza estadística.

PACS: 05.45.+b; 05.20.–y; 02.50.+s

1. Introduction and central results

Let us consider a system with states in a space X and with motion law given by the mapping $S: X \to X$, such that for each initial state x in X the successive states of our system at time $n = 1, 2, 3, \ldots$ are $S^1x = Sx, S^2x = S(Sx), S^3x = S(S^2x), \ldots$; hence, in its natural motion the system *deterministically* travels the orbit of states $\mathcal{O}_S(x) = \{x, Sx, S^2x, \ldots\}$. Similarly, for each ensemble of initial states $x_1, x_2, \ldots$ we obtain another ensemble of states $S^nx_1, S^nx_2, \ldots$ for each time $n = 1, 2, \ldots$; thus, the ensemble of initial states is transforming itself into another ensembles as time goes on, $n = 0, 1, 2, \ldots$; *i.e.*, the initial ensemble of states evolves in a natural form as time proceeds. The aim of this paper is to analyze this natural evolution for certain kind of initial ensembles of states in X.

The natural evolution of ensembles has been studied since several years ago, but these studies have been mainly concentrated on statistical mechanics, quantum mechanics and stochastic processes; however this analysis has been extended recently to a great variety of natural and mathematical systems. Among these dissimilar physical and mathematical systems (so far analyzed) we mention for future illustrations or references the following: the Maxwell-Boltzmann statistical mechanics [16] as well as the Einstein-Gibbs statistical mechanics [4,11]; a system of d independent and autonomous oscillators with incommensurable frequencies [3,21]; the thermodynamics out of equilibrium [22,29]; the chaotic advection [2]; in stochastic processes the weak convergence of probability measures [7,30] as well as the Markov chains [6,14,34]; several mappings on regions of $\mathbb{R}^n$ into itself with chaotic orbits like the logistic, the shifts, the baker's, etc. Many other systems can be found in Ref. [3,5,23,35].

In the framework of dynamical systems the above analysis has been made in the beautiful book by Lasota and Mackey [21], although it is aimed on those ensembles distributed by probability density functions with respect to *probability measures*. Using several ideas and extending some results in Ref. [21], we extend in the present article such analysis to ensembles distributed by probability density functions (PDF) with respect to *measures*.

With the aim of presenting a panoramic view of the content of this paper we will be more explicit. The fundamental elements of a dynamical system are: 1. A space X of all the states a system can take (X may be: the phase space in the Maxwell-Boltzman statistical mechanics, a surface of constant energy in the Einstein-Gibbs statistical mechanics; the phase space of a system of d oscillators; the space of positions in Aref's model of chaotic advection or the space of trajectories in a Markov chain). 2. The basic element through which the statistical regularities of the ensembles are originated is an extensive variable μ; we will consider those ensembles which are sequences of states distributed on X by probability density functions with respect to μ (μ-PDF); μ may be the length, area, volume,... of regions of $X = \mathbb{R}^m$ depending on $m = 1, 2, 3, \ldots$; or μ may be a probability measure, as in Markov chains or in weak converge of probability measures, *i.e.* μ measures something of each element in a collection Σ of subsets of X. 3. As it was explained at the beginning, a mapping S that maps points in X into points in X gives a motion law for discret times. A mapping like this may be found in a Hamiltonian system where X is a surface of constant energy and for each x in X, S is defined as $Sx = x(1)$, where $x(t)$ is the unique trajectory of the system such that $x(0) = x$ [18]; in the model of chaotic advection, where X is the space of positions, the mapping S is constructed through the Aref law of positions [2] in a similar way as before; in a shift S of Markov chain, where X is the space of all trajectories of the chain. X, Σ, μ and S are the fundamental elements of a dynamical system (DS) which will be denoted by the symbol (X, Σ, μ, S), and will be the framework of all the following.

The fundamental relations among the elements of our DS are conditioned by the assumption that for each ensemble $x_1, x_2, \ldots$ of states distributed on X by a μ-DPF f, its transformed ensemble $Sx_1, Sx_2, \ldots$ of states is distributed on X by a μ-DPF, Pf. And this fact is obtained from the hypotheses of Th. 2.1, which are very natural, as it can be seen in Sect. 4. Our extension to measures allow us to uncover (as a consequence of Th. 2.1) a peculiar region of X with the property that for almost each state x there is a time $n = n(x)$, such that all the states $S^n x, S^{n+1}x, \ldots$ are in that region; such region is called *absorbing attractor*, AA (but it is not strange attractor, Sect. 3). Therefore each ensemble in its natural evolution flows into the AA. Moreover, the orbit $\mathcal{O}_S(S^n x)$ happens to be quasi-ergodic in the AA. In the Aref's model of chaotic advection the AA is computed numerically; when some parameters are large (μ and t, see Ref. [2]) the final configuration of advected particles, *i. e.* the AA, is clearly observed in an extended region of the phase space. In a Markov shift constructed from a Markov chain having a finite number of states [6,34] with a unique irreducible closed set C, the resulting AA of this dynamical system is the set of all those trajectories $(s_0, s_1, \ldots)$ such that the s_n are in C for all but a finite number of values of n. In Sect. 3 is shown the AA of a relatively simple dynamical system.

The AA is the region where the unique stationary μ-DPF f_*(*i.e.* $Pf_* = f_*$) takes positive values, and it is denoted by $(f_* > 0)$. From the point of view of ensembles the equality $Pf_* = f_*$ means that for each ensemble $x_1^*, x_2^*, \ldots$ distributed on $(f_* > 0)$ by f_* each of its transformed ensembles $S^n x_1^*, S^n x_2^*, \ldots, n = 1, 2, \ldots$, is distributed on $(f_* > 0)$ by f_*; all those ensembles have the same statistical properties. Thus, each ensemble distributed by f_* is a microcanonical one.

An ensemble in its natural evolution flows into $(f_* > 0)$; in Sect. 4 it this behaviour is analyzed in terms of μ-PDFs. If an ensemble is distributed on X by the μ-PDF f, then the sequence of transformed ensembles has the sequences of μ-PDF's $Pf, P^2 f, \ldots$. In Th. 4.1, extension of part b) of Th. 4.4.1 in Ref. [21], that behaviour is characterized by the convergence of the sequence $P^n f$ to f_* *à la* weak Cesàro. That behaviour is very clear in the system of oscillators [21,25] and in the Ehrenfest chain [14]. (Apparently a similar behaviour happens in quantum billiards [36]).

We face from our framework the natural question that arises in statistical mechanics and in stochastic processes (Markov chains, weak convergence of probability measures, etc.): When does an ensemble of states in its natural evolution approach irreversibly to equilibrium? This is analyzed in Sect. 4 and it is obtained that the approach to equilibrium happens for every μ-PDF if and only if the DS (X, Σ, μ, S) is generalized mixing, Th. 4.1. Among the particular systems where the approach to equilibrium happens are Einstein-Gibbs statistical mechanics [23,28], Markov chains with irreducible and aperiodic probability transition matrices [6,14,34]. In the Aref's model it is observed experimentally. For several of the popular mappings such as the logistic, the unidirectional shifts, the baker's, etc. this approach to equilibrium also results [3,21,22].

A property of the generalized mixing DSs that is physically intuitive is *dispersivity*. This property is derived in Th. 5.2 and its corollary (enlargement of that for metric spaces of Erber *et al.* [13]): let a region A with $\mu(A) > 0$; then, under iterated applications of S the points in A tend to scatter on $(f_* > 0)$. A disturbing consequence of this property is the unpredictability in coarse-grained sense: from the knowledge that the system starts its motion on some initial conditions in A, no matter how small $\mu(A) > 0$ is, it is impossible to predict in what subregion of $(f_* > 0)$ will the state of the system be at each time n after certain time $n_0 = n_0(A)$. However, we are not helpless to face that randomness because a law of large numbers holds: if the system starts its motion from A N times and

$N_n(B)$ is the number of times that the system is in region B n time units latter, $n \geq n_0$, then $N_n(B) \longrightarrow_{n\to\infty} \int_B f_* \, d\mu$; and therefore, if N is large, then $N_n(B) \approx N \int_B f_* \, d\mu$.

2. Absorbing attractors

Our fundamental mathematical tools will be a measure space (X, Σ, μ) and a motion law given by a mapping $S: X \to X$; these together constitute a dynamical system, which will be denoted by (X, Σ, μ, S) or shortly by (μ, S). When that measure space is σ-finite and that mapping is Σ-measurable and μ-nonsingular (symbols and definitions are in the Appendix or in Ref. [21]) our dynamical system will be called fundamental, and shortly written by FDS. $P = P(\mu, S)$ will be the Frobenius-Perron operator of the FDS (μ, S). A in Σ is S-invariant if $S^{-1}A \underset{\mu}{=} A$. A dynamical system will be called G. ergodic (G. = generalized) if A being S-invariant, then either $A \underset{\mu}{=} \phi$ or $A \underset{\mu}{=} X$. With all these elements we have the following slight enlargement of Th. 4.4.1 in Ref. [21] which will be fundamental from now on.

Theorem 2.1. [27]. *Let (X, Σ, μ, S) be a FDS with at least one P-stationary density. Then, (μ, S) is G. ergodic $\Leftrightarrow$ P has a unique stationary density.*

Some definitions are needed previously to the statement of another fundamental theorem: for a real function f defined on X the symbol $(f > 0)$ will mean the set $\{x \in X; f(x) > 0\}$; with the unique P-stationary density function f_* we construct the probability measure μ_* defined on Σ as $\mu_*(B) = \int_B f_* \, d\mu$ for each B in Σ; as μ_* is absolutely continuous with respect to μ, we can write it shortly as $d\mu_* = f_* \, d\mu$ and it is the unique μ-absolutely continuous probability measure.

Theorem 2.2. *Let (X, Σ, μ, S) be a FDS with a unique P-stationary density function f_*. Let B in Σ such that $B \subset (f_* > 0)$ and $\mu_*(B) > 0$. Then the following statements are true:*

i) $X \underset{\mu}{=} \bigcap_{n\geq 1} \bigcup_{k\geq n} S^{-k}B$, *or with the usual probability notation,*

$$X \underset{\mu}{=} (S^{-n}B, \ i.o.);$$

ii) or shortly

$$X \underset{\mu}{=} \bigcup_{k\geq n} S^{-k}B$$

for every integer $n \geq 0$.

A first consequence of Th. 2.2 is the following: since $(f > 0) \subset S^{-1}(Pf > 0)$ for each f nonnegative μ-a.s. in $L_1(\mu)$, [21], we have $(f_* > 0) \subset S^{-1}(f_* > 0)$; furthermore

$$(f_* > 0) \subset S^{-1}(f_* > 0) \subset \ldots \subset S^{-n}(f_* > 0) \subset \ldots$$

Now, since $\mu_*(f_* > 0) = \int_{(f_*>0)} f_* \, d\mu = 1$, applying Th. 2.2 we obtain that $X \underset{\mu}{=} \bigcup_{k\geq 0} S^{-k}(f_* > 0)$. This expression means that for μ-a.e. x in X, $x \in \bigcup_{k\geq 0} S^{-k}(f_* > 0)$, and it also means that there exists an n such that $x \in S^{-n}(f_* > 0)$ or $S^n x \in (f_* > 0)$. This last belonging means that the orbit $\mathcal{O}_S(x)$ has entered the set $(f_* > 0)$; and $S^n x, S^{n+1}x, \ldots$ belong to $(f_* > 0)$ because of the previous chain of contentions. And if state y is in $(f_* > 0)$ then, $\mathcal{O}_S(y)$ never goes out of $(f_* > 0)$; that is, $\mathcal{O}_S(y) \subset (f_* > 0)$. All these properties are the reasons for calling the set $(f_* > 0)$ *absorbing attractor.*

Another relevant consequence of Th. 2.2 is that each orbit that enters the absorbing attractor visits each subset B of $(f_* > 0)$ an infinite number of times if $1 > \mu_*(B) > 0$; this is a consequence of applying Th. 2.2 to the sets B and B^c, since $1 > \mu_*(B^c) > 0$ too; since $B \underset{\mu}{=} \bigcap_{n\geq 0} \bigcup_{k\geq n} BS^{-k}B$ and $B \underset{\mu}{=} \bigcap_{n\geq 0} \bigcup_{k\geq n} BS^{-k}B^c$, for μ-a.e. x in $B, S^n x \in B$ for an infinite number of values of n, and $S^m x \in B^c$ for an infinite number of values of m, then the orbit $\mathcal{O}_S(x)$ enters and leaves an infinite number of times the set B. Then the set B is an attractor but not an absorbing one.

Let us note that this last consequence, considered from another point of view is just Poincare's recurrence theorem in the subspace $(f_* > 0)$ with measure μ_* and mapping S: the set of all those points belonging to B (with $\mu_*(B) > 0$) that always return to B after a finite number of applications of S, $\bigcap_{n\geq 0} \bigcup_{k\geq n} BS^{-k}B$, is equal to B, modulo μ_*.

The proof of Th. 2.2 is quite simple: $\bigcap_{n\geq 0} \bigcup_{k\geq n} S^{-k}B$ is an S-invariant set; since $\mu(B) > 0$ (if $\mu(B)$ is zero, then $\mu_*(B)$ would be zero since $\mu_* \ll \mu$), then

$$\mu(\bigcap_{n\geq 0} \bigcup_{k\geq n} S^{-k}B) > 0;$$

but the pair (μ, S) is G. ergodic by Th. 2.1, so that

$$X \underset{\mu}{=} \bigcap_{n\geq 0} \bigcup_{k\geq n} S^{-k}B.$$

Moreover, since

$$\bigcup_{k\geq m} S^{-k}B \supset \bigcap_{n\geq 0} \bigcup_{k\geq n} S^{-k}B$$

for each m, then

$$X \underset{\mu}{=} \bigcup_{k\geq m} S^{-k}B.$$

3. STATISTICAL REGULARITIES OF THE ORBITS' ELEMENTS

In order to establish the statistical regularities of the elements of almost each orbit, in Th. 2.2's framework, we must show that the mapping S preserves probability measure

μ_* and it is ergodic, *i.e.*, that the dynamical system (X, Σ, μ_*, S) is ergodic. (See the Appendix for a proof). Since $\mu_*(X - ((f_* > 0)) = 0$, the dynamical system $((f_* > 0), \Sigma(f_* > 0), \mu_*, S)$ (that in fact is the same as the one before) is ergodic, too.

We have seen that for each subset B of $(f_* > 0)$ with $0 < \mu_*(B) < 1$, for μ-a.e. x in X,the orbit $\mathcal{O}_S(x)$ enters and leaves the set B an infinite number of times. Moreover, for each set B in Σ, $\frac{1}{n}\sum_{k=0}^{n-1}\chi_B(S^k x) \underset{n}{\longrightarrow} \mu_*(B)$ follows for μ-a.e. x in X, as a consequence of an enlargement of a part of Birkhoff's ergodic theorem that will be seen in the proof of the following theorem. But the natural question is: if a system begins its motion on initial conditions x, will it visit *each* set B an infinite number of times and its asymptotic fraction of time of visits to B will be $\mu_*(B)$? It must be noted that Birkhoff's ergodic theorem says that *given* a set B, then $\frac{1}{n}\sum_{k=0}^{n-1}\chi_B(S^k x) \underset{n}{\longrightarrow} \mu_*(B)$ for μ_*-a.e. x in $(f_* > 0)$; *i.e.* that the set of points x such that this limit is true depends on the set B; that is to say: if we put $\tilde{B} = \{x \in (f_* > 0); \frac{1}{n}\sum_{k=0}^{n-1}\chi_B(S^k x) \underset{n}{\longrightarrow} \mu_*(B)\}$, then $\tilde{B} \underset{\mu_*}{=} (f_* > 0)$. To these questions there is an answer when X is a certain kind of manifold in $\mathbf{R}^n$, [8], but we only have a partial answer in Th. 2.2's framework:

Theorem 3.1. *Let (X, Σ, μ, S) be a FDS with a unique P-stationary density function f_* (then (μ, S) is G. ergodic). If the σ-field $\Sigma(f_* > 0)$ is separable (that is, if there exists a family of subsets $\beta = \{B_1, B_2, \ldots\}$ of Σ such that $\sigma(\beta) = \Sigma(f_* > 0)$), then for μ- a.e. x in X*

$$\frac{1}{n}\sum_{k=0}^{n-1}\chi_{B_i}(S^k x) \underset{n\to\infty}{\longrightarrow} \mu_*(B_i) = \int_{B_i} f_*\,d\mu,$$

for each B_i in β.

(See the Appendix for a proof).

It would certainly be very useful to enlarge this result to every set in Σ, but even the above version allows us to conclude: for a system with a motion law S in a space X (with the above characteristics) and μ-almost every initial conditions x in X,

i) the system's orbit $\mathcal{O}_S(x)$ will visit each set B_i an infinite number of times;

ii) the system's orbit $\mathcal{O}_S(x)$ will visit each B_i the asymptotic proportion of time $\mu_*(B_i)$, and

iii) the statistical distribution of the elements of almost each orbit is given by the probability measure μ_* on the family $\beta = \{B_1, B_2, \ldots\}$; that is, the orbit's elements are distributed on the absorbing attractor $(f_* > 0)$ with a probability law μ_* on β.

As an illustrative application, let the measure space be $([0,2], \mathcal{B}[0,2], \ell)$ where $\mathcal{B}[0,2]$ is the σ-field of Borel sets in the interval $[0,2]$ and ℓ is the Lebesgue measure defined on that σ-field; consider the mapping $S\colon [0,2] \to [0,2]$ such that

$$S(x) = \begin{cases} 4x(1-x) & \text{if } 0 < x \le 1 \\ -2x+4 & \text{if } 1 < x \le 2; \end{cases}$$

the measure space and S satisfy the Th. 2.2's hypotheses and the $P(\ell, S)$-stationary density function results to be

$$f_*(x) = \begin{cases} \dfrac{1}{\pi\sqrt{x(1-x)}} & \text{if } 0 \leq x \leq 1 \\ 0 & \text{if } 0 < x \leq 2; \end{cases}$$

the absorbing attractor results to be $(f_* > 0) = [0,1]$, where the mapping S is mixing with respect to measure

$$\mu_*(dx) = \frac{dx}{\pi\sqrt{x(1-x)}};$$

by Th. 3.1's conclusion the statistical distribution of the elements of $\mathcal{O}_S(x)$, for ℓ-a.e. x in $[0,2]$ is given by

$$\frac{1}{n}\sum_{k=0}^{n-1} \chi_{B_i}(S^k x) \xrightarrow[n]{} \int_{B_i \cap [0,1]} \frac{dx}{\pi\sqrt{x(1-x)}},$$

where the B_is are the intervals with rational extremes contained in the interval $[0,1]$. We have used the space $X = [0,2]$ to remark the role of $[0,1]$ as the absorbing attractor. If $[0,1]$ would be taken as the space X, that role would not be apparent.

On the other hand, the dynamical system (μ_*, S) can be mixing, exact, Kolmogorov or Bernoulli but the chaotic characteristics arise when (μ_*, S) is mixing (since when this happens the orbits have a certain kind of sensitive dependence on the initial conditions produced by a kind of dispersivity [13] as it will be seen in Sect. 4); when (μ_*, S) is at least mixing, absorbing attractors and strange attractors [5,8,10,31,33] may have the same properties from a randomness point of view, but not from a topological one. In the case of an absorbing attractor the topological properties are irrelevant being important only the measurable properties; but in strange attractors the topological properties are essential. On the other hand, if the space X is a manifold in $\mathbf{R}^n$, the absorbing attractor $(f_* > 0)$ never has a dimension less than that of X (if $(f_* > 0)$ is contained in a manifold of dimension less than that of X, then we have that $\mu(f_* > 0) = 0$, and this would be contrary to the fact that f_* is a $\mu-$density function); however, the dimension of a strange attractor is less than that of the manifold where it is immersed.

4. EVOLUTION OF ENSEMBLES

In the framework of a FDS(X, Σ, μ, S), whose Frobenius-Perron operator P has a unique stationary density function f_*, we are going to analyze how the ensembles obtained by iterated applications of S to an initial ensemble with μ-probability density function evolve. Let us consider an ensemble of state $x_1, x_2, \ldots$ in X distributed by a μ-probability density function f (by this we mean that the asymptotic fraction of members of the ensemble in every region B of space X is $\int_B f\,d\mu = \mu_f(B)$, for B in Σ, [22,28]; that is, $\frac{1}{N}\sum_{i=1}^{N}\chi_B(x_i) \underset{N\to\infty}{\longrightarrow} \int_B f\,d\mu = \mu_f(B)$).

Consider now the transformed ensemble of states $Sx_1, Sx_2, \ldots$, in X; as $\chi_B(Sx) = \chi_{S^{-1}B}(x)$ for every set B, we will have that

$$\frac{1}{N}\sum_{i=1}^{N}\chi_B(Sx_i) = \frac{1}{N}\sum_{i=1}^{N}\chi_{S^{-1}B}(x_i);$$

since the states $x_i, x_2, \ldots$ are distributed by the μ-probability density function f, we will have

$$\frac{1}{N}\sum_{i=1}^{N}\chi_B(Sx_i) = \frac{1}{N}\sum_{i=1}^{N}\chi_{S^{-1}B}(x_i) \underset{N\to\infty}{\longrightarrow} \int_{S^{-1}B} f\,d\mu;$$

and by the Frobenius-Perron operator P definition, $\int_{S^{-1}B} f\,d\mu = \int_B Pf\,d\mu$ results; thus,

$$\frac{1}{N}\sum_{i=1}^{N}\chi_B(Sx_i) \underset{N\to\infty}{\longrightarrow} \int_B Pf\,d\mu$$

is obtained. That is, the ensemble of states $Sx_1, Sx_2, ..$ has a μ-probability density function and this is Pf. (It must be noted that the probability measure $\mu_f S^{-1}$ has Pf as its μ-probability density function; that is, $\mu_f S^{-1}(B) = \int_B Pf\,d\mu$). In the same way it is evident that the ensemble $S^n x_1, S^n x_2, \ldots$ has $P^n f$ as its μ-probability density function, for $n = 0, 1, 2, \ldots$ (and the probability measure $\mu_f S^{-n}$ has $P^n f$ as its μ-probability density function; that is, $\mu_f S^{-n}(B) = \int_B P^n f\,d\mu$) [21].

A first feeling is that the ensemble $S^n x_1, S^n x_2, \ldots$ evolves toward $(f_* > 0)$ when n is increasing, because the action of the absorbing attractor $(f_* > 0)$ on almost every orbit, is a fact that is fully confirmed using Th. 2.2: we have that $S^{-n}(f_* > 0) \nearrow_n \bigcup_{k=0}^{\infty} S^{-k}(f_* > 0) \underset{\mu}{=} X$, and since μ_f is a probability measure we have $\mu_f S^{-n}(f_* > 0) \nearrow_n \mu_f(X) = 1$; but as $\int_{(f_*>0)} P^n f\,d\mu = \mu_f S^{-n}(f_* > 0) \nearrow_n 1$, the density function $P^n f$ is going to concentrate on the absorbing attractor $(f_* > 0)$ when n increases; but generally $(P^n f > 0)$ will not resemble to $(f_* > 0)$ as n increases, as the example below shows.

And, toward what and how $P^n f$ evolves when n increases?, does it converge to something? and if so, what kind of convergence is it? Let us make an ensemble evolution analysis: for the ensemble $x_1, x_2, \ldots$ with the μ-probability density function f, for each $i = 1, 2, \ldots$, we have $\frac{1}{n}\sum_{k=0}^{n-1}\chi_B(S^k x_i) \underset{n}{\longrightarrow} \mu_*(B)$ for each set B in the family β of Th. 3.1 (the points $x_1, x_2, \ldots$ can not be concentrated in a μ-measure zero set since $\mu(f > 0) > 0$); therefore we take those points in a set $X' \underset{\mu}{=} X$, the set of points for which the limit of

Th. 3.1 is true). For a set B in β and $x_1, \ldots, x_N$,

$$\frac{1}{M}\sum_{k=0}^{M-1}\chi_B(S^k x_1) \approx \mu_*(B)$$

$$\vdots$$

$$\frac{1}{M}\sum_{k=0}^{M-1}\chi_B(S^k x_N) \approx \mu_*(B)$$

as much as we like whenever M is large enough; adding these expressions term by term,

$$\frac{1}{N}\left(\frac{1}{M}\sum_{k=0}^{M-1}\chi_B(S^k x_1) + \cdots + \frac{1}{M}\sum_{k=0}^{M-1}\chi_B(S^k x_N)\right) \approx \mu_*(B)$$

is obtained; and this expression can be arranged as

$$\frac{1}{M}\sum_{k=0}^{M-1}\frac{\chi_B(S^k x_1) + \cdots + \chi_B(S^k x_N)}{N} \approx \mu_*(B);$$

and if N is large, then

$$\frac{\chi_B(S^k x_1) + \ldots + \chi_B(S^k x_N)}{N} \approx \int_B P^k f \, d\mu;$$

therefore, replacing this last approximation in the above approximation leads to

$$\frac{1}{M}\sum_{k=0}^{M-1}\int_B P^k f \, d\mu \approx \mu_*(B) = \int_B f_* \, d\mu.$$

So, we have arrived in a very natural way to the conclusion that for any μ-probability density function f the sequence $f, Pf, P^2 f, \ldots$ evolves to the density function f_* *à la* weak Cesàro; the apparently unnatural convergence *à la* weak Cesàro results in a very natural kind of convergence in a FDS with a unique absorbing attractor.

This conclusion answers the above questions; moreover, any such sequence $P^n f, n = 0, 1, 2, \ldots$ neither converges in a fine-grained (pointwise), nor in a coarse-grained (weakly), nor in a strong sense to anything. All this is confirmed by particular ergodic dynamical systems as the Ehrenfest Chain [15,36] and the quasi periodical motion. Let us consider the dynamical system $([0,1] \times [0,1], \mathcal{B}([0,1] \times [0,1]), \ell^{(2)}, S)$ where $\mathcal{B}([0,1] \times [0,1])$ is the σ-field of Borel sets of space $X = [0,1] \times [0,1]), \ell^{(2)}$ is the Lebesgue measure and the mapping S is the quasi periodic motion $S(x,y) = (\sqrt{2}, +x, \sqrt{3} + y)$ (Mod 1); this is essentially the motion of two independent and autonomous harmonic oscillators with angular velocities $\sqrt{2}$ and $\sqrt{3}$. In this case $f_* = 1$ and $(f_* > 0) = [0,1] \times [0,1]$; *i.e.*, $\ell^{(2)}$ is the unique S-invariant probability measure that is $\ell^{(2)}$-absolutely continuous. For the density function $\frac{\chi_A}{\ell^{(2)}(A)}$, where $A = [0,a] \times [0,a]$ with $0 < a < 1$, the sequence of densities $P^n\left(\frac{\chi_A}{\ell^{(2)}(A)}\right)$, $n = 0,1,2,\ldots$ has the sequence of sets $\left(P^n\left(\frac{\chi_A}{\ell^{(2)}(A)}\right) > 0\right)$, $n = 0,1,2,\ldots$, that seems to be bouncing on the walls of the square $[0,1] \times [0,1]$, [21].

In our framework of a FDS, the following theorem (enlargement of a) part of Th. 4.4.1 in Ref. [22]) gives some causes and consequences of the weak Cesàro convergence.

Theorem 4.1. *Let (X, Σ, μ, S) be a FDS. The following statements are equivalent:*

a) f_* is the unique P-stationary density function (therefore the dynamical system (μ, S) is G. ergodic);

b) $\mu_*: \Sigma \to [0,1]$, where $d\mu_* = f_*\, d\mu$, is the unique μ-absolutely continuous probability measure;

c) there exists a μ-absolutely continuous probability measure $\mu_*: \Sigma \to [0,1]$ with μ-density function f_*, such that

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu(AS^{-k}B) \underset{n}{\longrightarrow} \mu(A)\mu_*(B)$$

for all sets A and B in Σ with $\mu(A) < +\infty$;

d) there exists a μ-density function f_* such that for every μ-density function f

$$\frac{1}{n}\sum_{k=0}^{n-1} \int_B P^k f\, d\mu \underset{n}{\longrightarrow} \int_B f_*\, d\mu = \mu_*(B),$$

for every set B in Σ; i.e. the sequence of densities $P^n f, n = 0,1,2,\ldots$ converges to f_* à la weak Cesàro.

Moreover, any of these statements implies the dynamical system (μ_*, S) is ergodic.

(See the Appendix for a proof.)

The preceding results and the example give us a statistical-geometric idea about the natural evolution of ensembles of systems when there is only one absorbing attractor; first of all we know that as time proceeds, n increases, $P^n f$ tends to concentrate on the absorbing attractor $(f_* > 0)$, a fact that makes corresponding ensembles $S^n x_1, S^n x_2, \ldots$ go into the attractor $(f_* > 0)$, like a school of fish goes into a bay; once the ensemble $S^n x_1, S^n x_2, \ldots$, or part of it, is in the absorbing attractor it will travel from one place to another but always within $(f_* > 0)$, sometimes concentrating, sometimes extending and diluting in some zones, always varying in concentration as time goes by: $\int_B P^n f\, d\mu$ fluctuates when n increases, for each B zone of $(f_* > 0)$. And while this erratic tour proceeds, the ensembles $S^n x_1^*, S^n x_2^*, \ldots$, where $x_1^*, x_2^*, \ldots$ is an ensemble of the P-stationary density function f_*, are in statistical equilibrium: although for each fixed i, $S^n x_i^*$ is moving over all $(f_* > 0)$ because the ergodicity of S, and $S^n x_1^*, S^n x_2^*, \ldots$ is very different of $x_1^*, x_2^*, \ldots$, all these ensembles have the same statistical properties, *i.e.* the asymptotic proportion of ensemble elements $S^n x_1^*, S^n x_2^*, \ldots$ that are in B is the same for each time $n = 0, 1, 2, \ldots$; that is,

$$\frac{\chi_B S^n x_1^* + \cdots + \chi_B S^n x_N^*}{N} \xrightarrow[N]{} \int_B P^n f_* \, d\mu = \int_B f_* \, d\mu \xleftarrow[N]{} \frac{\chi_B x_1^* + \cdots + \chi_B x_N^*}{N},$$

for each B in Σ. The statistical properties of $S^n x_1^*, S^n x_2^*, \ldots$ do not change as time proceeds.

5. Irreversible evolution toward statistical equilibrium

As it can be seen, the preceding conditions assure that a system starting from almost each nonequilibrium state evolves towards its equilibrium in $(f_* > 0)$ as time proceeds, which is sufficient for the current value of a macroscopic observable to approach and stay very close to an equilibrium value. But, as noted above, those conditions do not assure that every ensemble with a μ-density function evolves to the stationary ensemble as time elapses, a situation that makes the dynamical functions of ensembles (as entropy [25]) will not approach their equilibrium values as time proceeds; hence, conditions are looked for to assure this kind of approach to stationarity [22,28], *i.e.* it is required that the ensembles $S^n x_1, S^n x_2, \ldots$ evolve irreversibly to the stationary ensemble $x_1^*, x_2^*, \ldots$ as n increases, in the sense that the statistical properties of $S^n x_1, S^n x_2, \ldots$ will be like those of $x_1^*, x_2^*, \ldots$, when n is sufficiently large, that is $\frac{\chi_B S^n x_1, + \cdots + \chi_B S^n x_N}{N} \approx \int_B f_* \, d\mu = \mu_*(B)$ when n is large; since $\frac{\chi_B S^n x_1 + \ldots + \chi_B S^n x_N}{N} \approx \int_B P^n f \, d\mu$, then $\int_B P^n f \, d\mu \approx \int_B f_* \, d\mu$ must hold. In precise terms, $\int_B P^n f \, d\mu \xrightarrow[n]{} \int_B f_* \, d\mu$ is required for each μ-probability density function f and each B in Σ.

For the time being the only answer we can give is similar to that wanted; therefore, only some consequences of that kind of evolution will be analyzed, and for this the following theorem is central:

Theorem 5.1. *Let (X, Σ, μ, S) be a FDS and f_* a μ-density function. The following statements are equivalent:*

a) *$\mu(AS^{-n}B) \xrightarrow[n]{} \mu(A)\, \mu_*(B)$ for all sets A and B in Σ with $\mu(A) < +\infty$, where $d\mu_* = f_* \, d\mu$;*

b) *for every μ-density function f, $P^n f \xrightarrow[n]{} f_*$ weakly; that is, $\int_B P^n f \, d\mu \xrightarrow[n]{} \int_B f_* \, d\mu = \mu_*(B)$ for each B in Σ.*

Moreover, each statement implies that f_* is the unique P-stationary density function (or, equivalently, that μ_* is the unique μ-absolutely continuous probability measure; therefore (μ, S) is G. ergodic) and the dynamical system (μ_*, S) is mixing.

(See the Appendix for a proof.)

A FDS (X, Σ, μ, S) with a μ-density function f_* is called *generalized mixing* (GM) if $\mu(AS^{-n}B) \underset{n\to\infty}{\longrightarrow} \mu(A)\mu_*(B)$ holds for every pair of sets A, B in Σ, with $\mu(A) < +\infty$. This kind of FDS generalizes the usual mixing dynamical systems: let A be a region of X, in Σ, with $0 < \mu(A) < +\infty$. If $a_1, a_2, \ldots,$ is an ensemble of states uniformly distributed over A, *i.e.* it has the μ-density function $\frac{\chi_A}{\mu(A)}$, then the chain of limits

$$\frac{1}{N}\sum_{i=1}^{N}\chi_B(S^n a_i) \underset{N\to\infty}{\longrightarrow} \int_B P^n\left(\frac{\chi_A}{\mu(A)}\right) d\mu \underset{n\to\infty}{\longrightarrow} \int_B f_*\, d_\mu = \mu_*(B)$$

holds for every B in Σ. Let us suppose the vessel $(f_* > 0)$ is filled with certain solvent and $a_1, a_2, \ldots, a_N$ are the molecules of the solute A, with N very large; then, for any region B of the vessel $(f_* > 0)$, $\sum_{i=1}^{N}\chi_B(S^n a_i)$ is the number of molecules of the solute A in B at time n. Then, using the above chain of limits, for time n sufficiently large, the number of molecules of A in the region B is approximately proportional to its "volume" $\mu_*(B)$, that is $\sum_{i=1}^{N}\chi_B(S^n a_i) \approx N\mu_*(B)$.

Now, a property of the GMFDS will be obtained because its central role in the following; this property is present in a great variety of physical and mathematical dynamical systems and it was firstly reported for some physical systems [12]. Consider a pair of subregions A, B of X, in Σ, with $\infty > \mu(A) > 0, B \subset (f_* > 0)$ and $0 < \mu_*(B) < 1$. Physically the set $AS^{-n}B = \{a \in A; S^n a \in B\}$ can be interpreted as those initial conditions in A for which the states of the system are in region B at time n. The *dispersivity property*, that we want to present, can be roughly settled as: a time n_0 must exist, such that all the states $S^n a$, with a in A, are not concentrated on B for each $n \geq n_0$; in other words, there exists an n_o such that $AS^{-n}B \underset{\mu}{=} A$ can not hold for each $n \geq n_o$. From the ensemble evolution point of view, the dispersivity property is very clear because each ensemble with μ-density function evolves, as time goes on, toward the stationary ensemble scattered by all the absorbing attractor; however, that is not a conclusive argument and a formal proof of that property is needed. Then, let us suppose $AS^{-n}B \underset{\mu}{=} A$ for an infinite number of n. We know that $\mu(AS^{-n}B) \underset{n}{\longrightarrow} \mu(A)\mu_*(B)$; thus, by the foregoing supposing, $\mu(A)\mu_i(B) = \mu(A)$; consequently, $\mu_*(B) = 1$, contrary to the assumption $0 < \mu_*(B) < 1$. The above property holds if we put B^c instead B, and it is proved in a similar way. We can summarize all this in the

Theorem 5.2. (GNFDS's dispersivity property). *Let (X, Σ, μ, S) be a GMFDS. If A, B are in Σ, with $0 < \mu(A) < +\infty, B \subset (f_* > 0)$ and $0 < \mu_*(B) < 1$ then there exists an n_o such that $0 < \mu(AS^{-n}B) < \mu(A)$ and $0 < \mu(AS^{-n}B^c) < \mu(A)$ for every $n \geq n_0$.*

This dispersivity property is a generalized version of that found out by Erber *et al.* for dynamical systems with S invertible and mixing with respect to a nonsingular and pervasive probability measure on a metric space [13]. However, such original version holds for a broader class of mixing dynamical systems as has been proved above (the unidirectional Bernoulli shifts, the logistic mapping over the unit interval and many more important dynamical systems cannot be considered by those original restrictions merely because their mappings are not invertible).

For our last result we will need the following corollary, whose intuitive meaning can be that each S^n granulates the set A over the absorbing attractor when n takes large values; in precise terms

Corollary to Th. 5.2. *If* $\{B_1, \ldots, B_M\}$ *is a partition of* $(f_* > 0)$ *with* B_i *in* Σ *and* $0 < \mu_*(B_i) < 1$ *for each* $i = 1, \ldots, M$*, then there exists an* n_o *such that* $0 < \mu(AS^{-n}B_i) < \mu(A)$ *for each* $i = 1, \ldots, M$*, whenever* $n \geq n_0$ *(In particular* $AS^{-n}B_i \neq \phi$ *when* $n \geq n_0$*).*

Finally, let us extract some consequences from the corollary. Consider a system with state space X and motion law S as in the corollary; we center our interest in a region A of X with a little measure $\mu(A) > 0$, from where the system starts its motion (A could represent the error arising from measurement impreciseness or numerical round offs). Moreover, to simplify let us suppose that there is an L such that $A \subset S^{-L}(f_* > 0)$, because Th. 2.1 (the forthcoming conclusions holds without that restriction); *i.e.*, if $a \in A$, then $S^L a \in (f_* > 0)$. If we take $n \geq \max\{n_0, L\}$, then $A = AS^{-n}(f_* > 0) = AS^{-n}B_1 + \cdots + AS^{-n}B_M$ with $0 < \mu(AS^{-n}B_i) < \mu(A)$, hence $AS^{-n}B_i \neq \phi$, for each $i = 1, \ldots, M$ because of the corollary. All that can be interpreted as: with the sole knowledge that the system starts its motion in some unknown state in A, the mere knowledge we can have is that the state of the system will be in some region B_i at time n. In other words: it is impossible to predict *in* what region B_i, for $i = 1, \ldots, M$, will the state of the system appear at time n, given that it started from initial conditions in A.

It will be advantageous to put the above conclusion in a probabilistic framework. Let us consider the experiment $\mathcal{E}_n$ = the system starts its motion from a state in A and it is observed n time units later, and the set of outcomes of $\mathcal{E}_n$ interesting to us is the sample space $\Omega = \{1, \ldots, M\}$ (the set of indexes of the regions $B_1, \ldots, B_M$). Our above issue is like having the pair $(\mathcal{E}_n\ \Omega)$: if in a trial of $\mathcal{E}_n$ the state of the system appears in region B_i, at time n, then the outcome i, in Ω, is obtained. In this framework, our above central conclusion becomes the essential property of the pair $(\mathcal{E}_n\ \Omega)$: the conditions under which the experiment is accomplished do not determine which of the possible outcomes, in Ω, will be obtained in each trial of $\mathcal{E}_n$; that is, the outcomes are random.

But it must be noticed that this random behaviour of GMFDS is in a regionwise, coarse-grained or macroscopic sense, and it is complementary to the deterministic behaviour, in pointwise, fine-grained or microscopic sense, a central assumption of DS (that is, given initial state x the system follows the determined orbit of state $\mathcal{O}_S(x)$). Although the orbits could be chaotic, or sensitive on initial conditions, right now, I am able to assure very little about that with the mathematical tools at our disposal in our framework; however, the randomness in regionwise, coarse-grained or macroscopic sense is inherent to chaotic orbits [9,31,32]. These two levels of apparently incompatible behaviour, or description [22], have been analyzed in several particular dynamical systems given by differential equations [1,3,5,15,17,19,20,23,24,26,29]; but, as we have seen above, these two levels of behaviour, or description, definitively hold for the wide class of the GMFDSs. (Moreover, the exact [21], Kolmogorov and Bernoulli dynamical systems are mixing dynamical systems [3,6,21,22]).

But the above randomness is not an irregular one. Let us consider an ensemble of initial conditions $a_1, a_2, \ldots$ in A; if we assume that the ensemble is uniformly distributed over

A, using the chain of limits just below Th. 5.1 we obtain that for each $j = 1, \ldots, M$, $\frac{1}{N}\sum_{i=1}^{N} \chi_{B_j}(S^n a_i) \approx \mu_i(B_j)$ as much as we like whenever $n \geq \max\{n_0, L\}$ and sufficiently large N. Therefore we can do statistical predictions: going back to our probability framework, if we make N trials of the experiment $\mathcal{E}_n$, then approximately $N\mu_*(B_j)$ times the state of the system is in B_j at time n. (The same result is obtained if we consider an ensemble with any other μ-density function concentrated on A).

And that is all we can predict about the future states of our system: after a certain time, we cannot do deterministic predictions; however, we can do predictions of a statistical nature.

APPENDIX

A. Definitions (more information in Ref. [21])

1) The indicator or characteristic function χ_A of set A is the function that takes value 1 in points in A and 0 value in points in A^c. If a mapping $S: X \to X$ is composed with indicator χ_A, where A is a subset of X, then $\chi_A \circ S = \chi_{S^{-1}A}$, *i.e.* $\chi_A(Sx) = \chi_{S^{-1}A}(x)$ for each x in X.

2) A measure space (X, Σ, μ) is σ-finite if there exists a denumerable partition $\{X_1, X_2, \ldots\}$ of space X, whose X_i are in Σ and $\mu(X_i) < +\infty$.

3) "μ-a.e. x in X" means almost every x in X with respect to measure μ, and "μ-a.s." means almost sure with respect to measure μ. $A \underset{\mu}{=} B$ means that $\mu(A\Delta B) = 0$, where $A\Delta B = (A - B) + (B - A)$ is the symmetric difference. With f and g measurable functions, $f \underset{\mu}{=} g$ means that $\mu(f \neq g) = 0$, where $(f \neq g) = \{x \in X; f(x) \neq g(x)\}$; that is, $f(x) = g(x)$ μ-a.e. x in X; $f \underset{\mu}{\geq}$ means that $\mu(f < 0) = 0$, where $(f < 0) = \{x \in X; f(x) < 0\}$, or that $(f \geq 0) \underset{\mu}{=} X$, or that $(f < 0) \underset{\mu}{=} \phi$.

4) For a Σ-measurable mapping $S: X \to X$ the measures $\mu S^{-n}: \Sigma \to [0, +\infty]$ can be constructed for each $n = 0, 1, 2, \ldots$, defined as $\mu S^{-n}(A) = \mu(S^{-n}A)$ where $S^{\circ}A = A$ for each set A in Σ. A measure ν is absolutely continuous with respect to μ, or ν is μ-absolutely continuous, when $\nu(A) = 0$ if $\mu(A) = 0$ for A in Σ, and it is denoted by $\nu \ll \mu$. S is said μ-nonsingular if $\mu S^{-1} << \mu$. By a μ-(probability) density function f we mean a real and measurable function $f: X \to \mathbb{R}$ such that $f \underset{\mu}{\geq}$ and $\int_X f\, d\mu = 1$. For each μ-density function f the probability measures $\mu_f S^{-n}: \Sigma \to [0, 1]$ can be constructed for each $n = 0, 1, 2, \ldots$ defined as $\mu_f S^{-n}(A) = \int_{S^{-n}A} f\, d\mu$ for each set A in Σ.

5) The Frobenius-Perron operator $P = P(\mu, S)$ of the dynamical system (μ, S) is the operator $P: L_1(X, \Sigma, \mu) \to L_1(X, \Sigma, \mu)$ that associates to each f in $L_1(\mu)$ a function Pf in $L_1(\mu)$ such that $\int_B Pf\, d\mu = \int_{S^{-1}B} f\, d\mu$ for each set B in Σ; such Pf exists and is μ-unique as a consequences of the Radon-Nikodym theorem applied to μ-nonsingular measure $\mu_f S^{-1}$ in the above measure space. If f is a μ-(probability) density function, then the probability measures $\mu_f S^{-n}, n = 0, 1, 2, \ldots$, are μ-absolutely continuous and their corresponding density functions are $P^n f, n = 0, 1, 2, ..$; that is $\mu_f S^{-n}(A) =$

$\int_A P^n f \, d\mu$ for each set A in Σ. A μ-density function is said P-stationary if $Pf = f$; then, a μ-density function f is P-stationary if and only if μ_f is S-invariant (or S preserves μ_f), that is $\mu_f S^{-1} = \mu_f$.

B. Proofs

1. In the framework of Th. 2.2. the pair (μ_*, S) is ergodic.

In the first place the pair (μ_*, S) is invariant, that is $\mu_* S^{-1} A = \mu_* A$ for each set A in Σ, since

$$\mu_* S^{-1}(A) = \int_{S^{-1}A} d\mu_* = \int_{S^{-1}A} f_* \, d\mu = \int_A P f_* \, d\mu = \int_A f_* d_\mu = \mu_*(A).$$

Now, if (μ_*, S) is not ergodic, then there exists a set A in Σ such that $S^{-1}A \underset{\mu}{=} A$ with $0 < \mu_*(A) < 1$. Let us see that the μ-density $\frac{\chi_A}{\mu_*(A)} f_*$ is a P-invariant density function: as $\chi_A \underset{\mu_*}{\equiv} \chi_{S^{-1}A}$ then

$$\frac{\chi_A \cdot f_*}{\mu_*(A)} \underset{\mu_*}{\equiv} \frac{\chi_{S^{-1}A} \cdot f_*}{\mu_*(A)}.$$

By the Frobenius-Perron operator,

$$\int_{S^{-1}B} \frac{\chi_{S^{-1}A}}{\mu_*(A)} f_* \, d\mu = \int_B P\left(\frac{\chi_{S^{-1}A}}{\mu_*(A)} f_*\right) d\mu.$$

On the other hand,

$$\int_{S^{-1}B} \frac{\chi_{S^{-1}A}}{\mu_*(A)} f_* \, d\mu = \int_{S^{-1}B} \frac{\chi_{A} \circ S}{\mu_*(A)} f_* \, d\mu = \int_B \frac{\chi_A}{\mu_*(A)} \, d\mu_* S^{-1}.$$

Since $\mu_* S^{-1} = \mu_*$, then

$$\int_{S^{-1}B} \frac{\chi_{S^{-1}A}}{\mu_*(A)} f_* \, d\mu = \int_B \frac{\chi_A}{\mu_*(A)} \, d\mu_* = \int_B \frac{\chi_A}{\mu_*(A)} f_* \, d\mu,$$

but

$$\int_B \frac{\chi_A}{\mu_*(A)} f_* d_\mu = \int_B \frac{\chi_{S^{-1}A}}{\mu_*(A)} f_* \, d\mu$$

therefore

$$\int_{S^{-1}B} \frac{\chi_{S^{-1}A}}{\mu_*(A)} f_* \, d\mu = \int_B \frac{\chi_{S^{-1}A}}{\mu_*(A)} f_* \, d\mu.$$

With all this we get that

$$\int_B P(\frac{\chi_{S^{-1}A}}{\mu_*(A)} f_*)\, d\mu = \int_B \frac{\chi_{S^{-1}A}}{\mu_*(A)} f_*\, d\mu;$$

since this equality holds for every B in Σ, then

$$P\left(\frac{\chi_{S^{-1}A}}{\mu_*(A)} f_*\right) = \frac{\chi_{S^{-1}A}}{\mu_*(A)} f_*,$$

which means that density

$$\frac{\chi_{S^{-1}A}}{\mu_*(A)} f_* \underset{\mu_*}{=} \frac{\chi_A}{\mu_*(A)} f_*$$

is another P-stationary density function, contrary to initial assumption that f_* was the unique P-stationary density function.

2. Th. 3.1. Let us first stablish a slight enlargement of Birkhoff's ergodic theorem:

Lemma. *With the hypothesis of Th. 2.1 and with f_* being the unique P-stationary density function, for each B in Σ, $\frac{1}{n}\sum_{k=0}^{n-1} \chi_B(S^k x) \underset{n}{\longrightarrow} \mu_*(B)$ follows for μ-a.e. x in X.*

Proof: since the pair (μ_*, S) is ergodic, for each B in Σ $\frac{1}{n}\sum_{k=0}^{n-1} \chi_B(S^k x) \underset{n}{\longrightarrow} \mu_*(B)$ follows for μ_*-a.e. x in X because of Birkhoff's ergodic theorem. Denoting by $\tilde{B}$ the set of points in X for which the foregoing limit holds, we will have that $\mu_*(\tilde{B}) = 1$; but $S^{-1}\tilde{B} = \tilde{B}$, then $\mu(\tilde{B}) = 0$ or $\mu(\tilde{B}^c) = 0$ since (μ, S) is G. ergodic. If $\mu(\tilde{B}) = 0$ is true, then $\mu_*(\tilde{B}) = 0$ because $\mu_* \ll \mu$; then it must be that $\mu(\tilde{B}^c) = 0$, that is $X \underset{\mu}{=} \tilde{B} = \{x; \frac{1}{n}\sum_{k=0}^{n-1} \chi_B(S^k x) \underset{n}{\longrightarrow} \mu_*(B)\}$.

Now the proof of Th. 3.1: since for every B_i in β we have $\tilde{B}_i \underset{\mu}{=} X$, it follows that $\bigcap_{i=1}^{\infty} \tilde{B}_i \underset{\mu}{=} X$, which means that for μ-a.e. point x in X we will have

$$\frac{1}{n}\sum_{k=0}^{n-1} \chi_{B_i}(S^k x) \underset{n}{\longrightarrow} \mu_*(B)$$

for every B_i in β.

3. Th. 4.1.

i) $a) \Leftrightarrow b)$ follows immediately.

ii) Let us suppose that $a)$ or $b)$ holds. Since P has a unique stationary density, using the preceding *Lemma* it follows that for each set B in Σ, $\frac{1}{n}\sum_{k=0}^{n-1} \chi_B(S^k x) \underset{n}{\longrightarrow} \mu_*(B)$ for μ-a.e x in X. Now, for each set A in Σ it follows that $F_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} \chi_A(x) \cdot \chi_B(S^k x) \underset{n}{\longrightarrow}$

$\chi_A(x) \cdot \mu_*(B)$ for μ-a.e. x in X. Since: a) $F_n(x) \leq \chi_A(x)$ for each x in X; b) χ_A is μ-integrable if $\mu(A) < +\infty$, and c) $F_n(x) \xrightarrow[n]{} \mu_*(B)\chi_A(x)$ for μ-a.e. x in X; applying Lebesgue's Dominated Convergence Theorem it follows that

$$\int_X F_n(x)\mu(dx) \xrightarrow[n]{} \int_X \mu_*(B)\chi_A(x)\mu(dx).$$

But

$$\int_X F_n(x)\mu(dx) = \int_X \frac{1}{n}\sum_{k=0}^{n-1} \chi_A(x) \cdot \chi_B(S^k x)\mu(dx)$$

$$= \frac{1}{n}\sum_{k=0}^{n-1} \int_X \chi_A \cdot (\chi_{B^\circ} S)\, d\mu = \frac{1}{n}\sum_{k=0}^{n-1} \mu(AS^{-k}B);$$

also

$$\int_X \mu_*(B)\chi_A(x)\mu(dx) = \mu_*(B)\mu(A).$$

Then, replacing these in the previous limit, we obtain

$$\frac{1}{n}\sum_{k=0}^{n-1} \mu(AS^{-k}B) \xrightarrow[n]{} \mu(A)\mu_*(B).$$

So, *a*) or *b*) $\Rightarrow$ *c*).

iii) Let us show that *c*) $\Rightarrow$ *d*). Furthermore, we prove that for each $f \underset{\mu}{\geq}$ in $L_1(\mu)$

$$\frac{1}{n}\sum_{k=0}^{n-1} \int_B P^k f\, d_\mu \xrightarrow[n]{} \mu_*(B) \int_X f\, d\mu,$$

and when f is a μ-density function the limit in *d*) follows.

1°. $f = \chi_A$ in $L_1(\mu)$: since $\int_B P^k(\chi_A)\, d\mu = \mu(AS^{-k}B)$, then

$$\frac{1}{n}\sum_{k=0}^{n-1} \int_B P^k(\chi_A)\, d\mu = \frac{1}{n}\sum_{k=0}^{n-1} \mu(AS^{-k}B) \xrightarrow[n]{} \mu_*(B)\mu(A) = \mu_*(B) \int_X \chi_A\, d\mu.$$

From now on, we will use the symbol $\mathbf{P}_n(g)$ to denote

$$\mathbf{P}_n(g) = \frac{1}{n}\sum_{k=0}^{n-1} \int_B P^k g\, d\mu;$$

2° let a simple function $f = \sum_{i=1}^{m} a_i \chi_{A_i}$ in $L_1(\mu)$; then

$$\mathbf{P}_n\left(\sum_{i=1}^{m} a_i\chi_{A_i}\right) = \sum_{i=1}^{m} a_i\mathbf{P}_n(\chi_{A_i}) \xrightarrow[n]{} \sum_{i=1}^{m} a_i\mu_*(B)\mu(A_i) = \mu_*(B)\int_X \sum_{i=1}^{m} a_i\chi_{A_i}\,d\mu,$$

as a consequence of 1° above.

3°. Let $f \underset{\overline{\mu}}{\geq}$ in $L_1(\mu)$. If $\{S_m\}$ is a sequence of simple functions, with each S_m in $L_1(\mu)$, such that $0 \leq S_m \nearrow_m f$ and $\int_X S_m\,d\mu \xrightarrow[m]{} \int_X f\,d\mu$, we will have the following:

$$\begin{aligned}\left|\mathbf{P}_n(f) - \mu_*(B)\int_X f\,d\mu\right| &\leq |\mathbf{P}_n(f) - \mathbf{P}_n(S_m)| + \left|\mathbf{P}_n(S_m) - \mu_*(B)\int_X f\,d\mu\right| \\ &\leq |\mathbf{P}_n(f) - \mathbf{P}_n(S_m)| + \left|\mathbf{P}_n(S_m) - \mu_*(B)\int_X S_m\,d\mu\right| \\ &\quad + \left|\mu_*(B)\int_X S_m d\mu - \mu_*(B)\int_X f\,d\mu\right|;\end{aligned}$$

but

$$\begin{aligned}|\mathbf{P}_n(f) - \mathbf{P}_n(S_m)| &\leq \frac{1}{n}\sum_{k=0}^{n-1}\left|\int_B P^k(f - S_m)\,d\mu\right| \\ &= \frac{1}{n}\sum_{k=0}^{n-1}\int_{S^{-1}B}(f - S_m)\,d\mu \leq \int_X (f - S_m)\,d\mu.\end{aligned}$$

Now, given $\varepsilon_1 > 0$ there exists a $m_o(\varepsilon_1)$ such that $\int_X (f - S_m)\,d\mu < \varepsilon_1$ for every $m \geq m_o(\varepsilon_1)$, and therefore $|\mathbf{P}_n(f) - \mathbf{P}_n(S_m)| < \varepsilon_1$ for every $m \geq m_0(\varepsilon_1)$.

By 2° above, given $\varepsilon_2 > 0$ there exists a $n_0(\varepsilon_2, m)$ such that $|\mathbf{P}_n(S_m) - \mu_*(B)\int_X S_m\,d\mu| < \varepsilon_2$ for every $n \geq n_0(\varepsilon_2, m)$, and

$$\left|\mu_*(B)\int_X S_m\,d\mu - \mu_*(B)\int_X f\,d\mu\right| = \mu_*(B)\left(\int_X f\,d\mu - \int_X S_m\,d\mu\right) < \mu_*(B)\varepsilon_1$$

for every $m \geq m_0(\varepsilon_1)$.

Now, taking $m \geq m_o(\varepsilon_1)$ and $n \geq n_0(\varepsilon_2, m)$ we get

$$\left|\mathbf{P}_n(f) - \mu_*(B)\int_X f\,d\mu\right| < \varepsilon_1 + \varepsilon_2 + \mu_*(B)\varepsilon_1,$$

which means that

$$\frac{1}{n}\sum_{k=0}^{n-1}\int_B P^k f\,d\mu \underset{n}{\longrightarrow} \int_B f_*\,d\mu \int_X f\,d\mu = \mu_*(B)\int_X f\,d\mu.$$

And if f is a μ-density function, then

$$\frac{1}{n}\sum_{k=0}^{n-1}\int_B P^k f\,d\mu \underset{n}{\longrightarrow} \int_B f_*\,d\mu = \mu_*(B).$$

With all this $c) \Rightarrow d)$ has been proved.

iv) $d) \Rightarrow c)$: if $f = \frac{\chi_A}{\mu(A)}$ with $0 < \mu(A) < +\infty$, then f is a μ-density function, and

$$\frac{1}{n}\sum_{k=0}^{n-1}\frac{\mu(AS^{-k}B)}{\mu(A)} = \frac{1}{n}\sum_{k=0}^{n-1}\int_B P^k\left(\frac{\chi_A}{\mu(A)}\right)d\mu \underset{n}{\longrightarrow} \mu_*(B)\int_X \frac{\chi_A}{\mu(A)}\,d\mu = \mu_*(B),$$

then

$$\frac{1}{n}\sum_{k=0}^{n-1}\mu(AS^{-k}B) \underset{n}{\longrightarrow} \mu(A)\mu_*(B)$$

for all sets A and B in Σ with $\mu(A) < +\infty$.

Thus, with iii) and iv), we have that $c) \Leftrightarrow d)$.

v) Let us see that $c) \Rightarrow a)$ or $b)$. If $\frac{1}{n}\sum_{k=0}^{n-1}\mu(AS^{-k}B) \underset{n}{\longrightarrow} \mu(A)\mu_*(B)$ for all sets A and B in Σ with $\mu(A) < +\infty$, then

I. $\mu_* S^{-1} = \mu_*$, *i.e.*, the pair (μ_*, S) is invariant. Since

$$\frac{1}{n}\sum_{k=0}^{n-1}\mu\big(AS^{-k}(S^{-1}B)\big) \underset{n}{\longrightarrow} \mu(A)\mu_*(S^{-1}B)$$

and

$$\frac{1}{n}\sum_{k=0}^{n-1}\mu(AS^{-k-1}B) = \frac{n+1}{n}\frac{1}{n+1}\sum_{k=0}^{n}\mu(AS^{-k}B) - \frac{\mu(AB)}{n} \underset{n}{\longrightarrow} \mu(A)\mu_*(B),$$

then $\mu_*(S^{-1}B) = \mu_*(B)$ for each B in Σ, that is $\mu_* S^{-1} = \mu_*$.

II. Now, since $d\mu_* = f_*\,d\mu$ is a S-invariant $\Leftrightarrow Pf_* = f_*$, it follows that our f_* is P-stationary density function.

III. If there would be another P-stationary density function g_* such that

$$\frac{1}{n}\sum_{k=0}^{n-1}\mu(AS^{-k}B) \xrightarrow[n]{} \mu(A)\nu_*(B),$$

where $d\nu_* = g_*\,d\mu$, we could have $\nu_*(B) = \mu_*(B)$ for each set B in Σ. But this last equality is $\int_B g_*\,d\mu = \int_B f_*\,d\mu$ for each set B in Σ, which in turn implies that $g_* \underset{\mu}{=} f_*$. Then, there is a unique P-stationary density function f_* or there is only one μ-absolutely continuous probability measure μ_*. With all this we have $c) \Rightarrow a)$ and $b)$.

To complete the proof of Th. 4.1 remember that if there is a unique P-stationary density function f_*, then the pair (μ_*, S) is ergodic.

4.Th. 5.1

i) $a) \Rightarrow b)$. The proof is similar to that given for Th. 1.3, but taking off the symbol $\frac{1}{n}\sum_{k=0}^{n=1}$, and getting $\int_B P^n f\,d\mu \xrightarrow[n]{} \int_B f_*\,d\mu \int_X f\,d\mu = \mu_*(B)\int_X f\,d\mu$ for each $f \underset{\mu}{\geq}$ in $L_1(\mu)$.

ii) $b) \Rightarrow a)$. If $f = \frac{\chi_A}{\mu(A)}$ with $0 < \mu(A) < +\infty$ then f is a μ-density function and therefore

$$\frac{\mu(AS^{-n}B)}{\mu(A)} = \int_B P^n\left(\frac{\chi_A}{\mu(A)}\right)d\mu \xrightarrow[n]{} \mu_*(B)\int_X \frac{\chi_A}{\mu(A)}\,d\mu = \mu_*(B);$$

then, $\mu(AS^{-n}B) \xrightarrow[n]{} \mu(A)\mu_*(B)$ for any sets A and B in Σ with $\mu(A) < +\infty$.

iii) The fact that f_* is the unique P-stationary density function or that μ_* is the unique μ-absolutely continuous probability measure follows in a similar way as in Th. 3.1 but taking off the symbol $\frac{1}{n}\sum_{k=0}^{n-1}$ or considering that weak convergence is a particular case of weak Cesàro convergence.

That (μ_*, S) is mixing follows from

$$\frac{\mu_*(AS^{-n}B)}{\mu_*(A)} = \int_{AS^{-n}B}\frac{f_*}{\mu_*(A)}\,d\mu = \int_B P^n\left(\frac{\chi_A}{\mu_*(A)}f_*\right)d\mu \xrightarrow[n]{} \int_B f_*\,d\mu = \mu_*(B),$$

because the function $\frac{\chi_A}{\mu_*(A)}f_*$ is a μ-density function when $0 < \mu(A) < +\infty$; then $\mu_*(AS^{-n}B) \xrightarrow[n]{} \mu_*(A)\mu_*(B)$ for all sets A and B in Σ. Then, we have proved Th. 5.1.

ACKNOWLEDGEMENTS

The autor is grateful to Professors D. Moreno and I. Campos (Departamento de Física de la Facultad de Ciencias de la UNAM) for the revision and opinions about this work. It is also a pleasure to aknowledge helpful discussions with them.

REFERENCES

1. V.M.A. Alekseev, *Math. U.S.S.R. Sbornik,* **5** (1968) 53; **6** (1968) 505; **7** (1969) 1.
2. H. Aref, *J. Fluid. Mech.* **143** (1984) 1.
3. V.I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, W.A. Benjamin, Inc. (1968).
4. A. Baracca and R. Rechtman, *Proceedings of the Second Escuela Mexicana de Física Estadística*, Eds. R. Peralta-Fabi and F. Ramos-Gómez, Soc. Mex. Fís. (1985).
5. P. Bergé, Y. Pomeau et Ch. Vidal, *L'ordre dans le chaos*, Hermann, Paris (1984).
6. P. Billingsley, *Ergodic Theory and Information*, J. Wiley and Sons (1965).
7. P. Billingsley, *Convergence of Probability Measures*, J. Wiley and Sons (1968).
8. R. Bowen and D. Ruelle, *Inventiones Math.* **29** (1975) 181.
9. R.L. Devaney, *An Introduction to Chaotic Dynamical Systems*, The Benjamin Cummings Publishing Co., Inc. (1986).
10. J.P. Eckmann and D. Ruelle, *Rev. Mod. Phys.* **57** (1985) 617.
11. P. and T. Ehrenfest, *The Conceptual Foundations of Statistical Approach in Mechanics*, Cornell University Press (1959).
12. T. Erber and A. Sklar, *Modern Developments in Thermodynamics*, ed. by Gal-Or. Jerusalem-New York: Israel University Press and J. Wiley and Sons (1973).
13. T. Erber, B. Schweizer and A. Sklar, *Commun. Math. Phys.* **29** (1973) 311.
14. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I, J. Wiley and Sons Inc. (1968).
15. J. Ford, *Phys. Today* (April, 1983) 40.
16. L. García-Colín S., *Rev. Mex. Fís.,* **28** (1982) 121.
17. J.B. Keller, *Amer. Math. Monthly*, (March, 1986) 191.
18. A.I. Khinchin, *Mathematical Foundations of Statistical Mechanics*, Dover Publications, Inc. (1949).
19. J. Laskar, *Nature* **338** (1989) 237.
20. J. Laskar et C. Froeschlé, *La Recherche* **22** (1991) 572.
21. A. Lasota and M.C. Mackey, *Probabilistic Properties of Deterministic Systems*, Cambridge University Press (1985).
22. J.L. Lebowitz and O. Penrose, *Phys. Today* (February, 1973) 23.
23. A.J. Lichtenberg and M.a. Lieberman, *Regular and Stochastic Motion*, Springer-Verlag (1983).
24. J. Lighthill, *Proc. R. Soc. Lond* **A 407** (1986) 35.
25. N.C. Mackey, *Rev. Mod. Phys.* **61** (1989) 181.
26. J. Moser, *Stable and Random Motions in Dynamical Systems*, Princenton University Press and University of Tokyo Press. Princenton, New Jersey (1973).
27. A. Nieva, *Stochastic Anal. and App.* **10** (1992) 313.
28. O. Penrose, *Foundations of Statistical Mechanics*, Pergamon Press (1970).
29. I. Prigogine et I. Stengers, *Entre le temps et l'éternité*, Libraire Arthème Fayard (1988).
30. Yu. v. Prohorov, *Theor. Prob. Appl.* **1** (1956) 157.
31. D. Ruelle, *Chaotic Evolution and Strange Attractors*, Cambridge University Press (1989).
32. D. Ruelle, *Elements of Differentiable Dynamics and Bifurcation Theory*, Academic Press, Inc. (1989).
33. D. Ruelle and F. Takens, *Commun. Math. Phys.* **20** (1971) 167.
34. P. Walters, *An. Introduction to Ergodic Theory*, Springer Verlag (1982).
35. G. Zaslavsky, *Chaos in Dynamic Systems*, Harwood Academic (1985).
36. G. Casati, *La Recherche,* **22** (1991) 600.