Journal of Statistical Physics, Vol. 47, Nos. 5/6, 1987

Dynamical Systems with Relaxation Time

Martin Macháček[1]

Received November 7, 1986

We define mathematically a class of dynamical systems that exhibit relaxation corresponding to that observed in physical systems, and then show that this class is identical with the class of K-mixing dynamical systems.

KEY WORDS: Dynamical systems; relaxation time; K-mixing; entropy.

Equilibrium statistical mechanics of Hamiltonian systems is based on the observation that after a certain time—called the relaxation time—some isolated Hamiltonian systems, such as a gas, come from any initial conditions to a state of statistical equilibrium. In this state we can, by measuring macroscopic observables, learn nothing more about the initial conditions of the system than the values of conserved quantities (for definiteness we shall speak of energy E, momentum P, and angular momentum M); this is equivalent to the probability distribution being given by the ergodic invariant measure corresponding to these values of E, P, and M, the microcanonical distribution. Hence the time average $T^{-1}\int_0^T f(X_t)\,dt$ of a function f of the microstate X_t of the system approaches its ensemble average with respect to the microcanonical distribution, $\int f(x)\,d\mu(x)$, if only T is of the order of several times the relaxation time. Reed and Simon[1] raise the question of what properties of the dynamical system are responsible for this phenomenon of relaxation. We shall see that the answer can be "K-mixing." I say "can be" because so far we have defined relaxation and the class of systems that possess it rather loosely; we shall give a physically reasonable definition and then show that the class of systems with relaxation time is the same as the class of K-mixing systems.

[1] Astronomical Institute of the Czechoslovak Academy of Sciences, 251 65 Ondřejov, Czechoslovakia.

0022-4715/87/0600-0949$05.00/0

Let Ω be the phase space of the system and x its points, the microstates; the trajectory of the system in Ω will be denoted by X_t. By measurements we determine the microstate of the system not completely, but by a rather small number, say M, of values of phase functions $g_m: \Omega \to \mathbb{R}$, $m = 1,..., M$. The equivalence between states indistinguishable from each other by these measurements induces a partition ξ of Ω into macrostates, $\xi = \{A_i\}$; each A_i consists of the points x for which $g_m(x)$ have the same value within the accuracy of the measurement. Let the dynamical group be $\{T_t\}$; hence $X_t = T_t X_0$. Define, as usual, $T_t^{-1}\xi = \{T_t^{-1}A_i\}$; measuring g at the time t, we find to which A_i the state $X_t = T_t X_0$ belongs, or to which set $T_t^{-1}A_i$ of the partition $T_t^{-1}\xi$ the state X_0 belongs; and measuring g at the times $t_1,..., t_n$, we find to which set of the partition $T_{t_1}^{-1}\xi T_{t_2}^{-1}\xi \cdots T_{t_n}^{-1}\xi$ the point X_0 actually belongs. (The product of partitions $\xi = \{A_i\}$ and $\eta = \{B_j\}$ is defined by $\xi\eta = \{A_i \cap B_j\}$.)

The partition corresponding to the phase functions $E(x)$, $P(x)$, and $M(x)$ is invariant against the evolution; $T_t X_0$ remains forever in the same set of this partition. The phenomenon of relaxation occurs within these sets, or, as we would normally say, within energy hypersurfaces. Therefore we may take one of these energy hypersurfaces for the phase space. We shall assume that after this restriction our dynamical system is strictly ergodic, or, in other words, that there is only one invariant measure (IM) concentrated on each energy hypersurface. For classical physical systems such as a gas there are no experimental facts that contradict this assumption.

Now we come to the amount of information[2] contained in communicating the result of measurement of the quantities g_m. If we have no prior knowledge of the state of the system, this amount is, as is well known,

$$H(\xi) = -\sum_i p_i \ln p_i, \qquad p_i = \mu(A_i), \quad A_i \in \xi$$

where μ is the unique IM. If we know in advance that the system is in a set B_j of another partition η, then the amount of information is

$$-\sum_i \mu(A_i \mid B_j) \ln \mu(A_i \mid B_j), \qquad \mu(A_i \mid B_j) = \mu(A_i \cap B_j)/\mu(B_j)$$

and its average is the conditional entropy

$$H(\xi \mid \eta) = -\sum_{ij} \mu(A_i \cap B_j) \ln \mu(A_i \mid B_j)$$

In particular, $H(\xi \mid T_t^{-1}\xi) \equiv H(\xi \mid \xi_{-t})$ is the mean amount of information contained in communicating the results of measuring $g(X_0)$ if the results of measuring $g(X_t)$ are known in advance, and $H(\xi \mid \xi_{-t_1}\xi_{-t_2}\cdots\xi_{-t_n})$ is the amount of new information contained in $g(X_0)$ provided $g(X_{t_1}),..., g(X_{t_n})$ are known. If we perform a series of measurements from the time t on, we can infer something about the state X_0 at the time 0, but the amount of information thus obtained is bounded by

$$H(\xi) - H(\xi \mid \xi_{-\infty}^{-t}) \tag{1}$$

where

$$H(\xi \mid \xi_{-\infty}^{-t}) = \inf_{\substack{t_1 \geq t \\ t_2 \geq t \\ \vdots}} H(\xi \mid \xi_{-t_1}\xi_{-t_2}\cdots)$$

If the dynamical system exhibits relaxation, then, as t grows, the knowledge of $g(X_{t_1})$, $g(X_{t_2})$,... contributes less and less to our ability to say something about $g(X_0)$, hence $H(\xi \mid \xi_{-\infty}^{-t})$ and $H(\xi)$, the amounts of information contained in $g(X_0)$ with and without prior knowledge of the values of g from the time t on, become almost equal. Therefore we could use the way in which (1) approaches zero as an indication of relaxation; and we could be more general, measuring g at the time 0 and another set of macrovariables $f_{m'}$, $m' = 1,..., M'$ (which corresponds to a partition η), from the time t on.

After these preliminaries, we define relaxation time (RT) as

$$\tau(\xi, \eta; \varepsilon) = \inf\{T \mid H(\xi) - H(\xi \mid \eta_{-\infty}^{-t}) < \varepsilon \text{ for all } t \geq T\} \tag{2}$$

This means that measuring f later than $\tau(\xi, \eta; \varepsilon)$, we cannot obtain more information than ε about $g(X_0)$. If $\tau(\xi, \eta; \varepsilon) < \infty$ for all $\varepsilon > 0$, we say that the system has RT with respect to the pair (ξ, η); this is equivalent to $\lim_{\tau \to \infty} \varepsilon(\xi, \eta; \tau) = 0$, where $\varepsilon(\xi, \eta; \tau)$ is the inverse function to $\tau(\xi, \eta; \varepsilon)$. Finally, if the system has RT with respect to any pair (ξ, η), where ξ is a nontrivial partition (that containing no elements of measure 1—the functions g_m must not be all constant μ-almost everywhere), then we drop the qualification and say that it has RT. The meaning of this property is that as t grows, it becomes impossible to retrodict the value of any function (except a constant one) of X_0, no matter what we measure at times later than t.

Now we can easily prove that the systems with RT are exactly the K-mixing systems. There are several characterizations of the K-mixing systems and one of them that suits our purposes is, for every two partitions ξ and η (ξ nontrivial), we have $H(\xi \mid \eta_{-\infty}^{-t}) \to H(\xi)$ as $t \to \infty$. This means precisely that the system has RT.

We have not defined the RT of a dynamical system, only the property of having RT. This is in accord with physical experience: RT is not a numerical quantity. However, for a given set (or pair of sets) of macroscopic observables, RT can be estimated to an order of magnitude; while this estimate may differ substantially when we go from one set of observables to another (e.g., from pressures to temperatures), it is relatively insensitive to improvements of measurement precision. The definition (2) of RT accounts for these empirical facts. As the uncertainty Δg_m of g_m is diminished, $\mu(A_i)$ varies roughly proportionally to $\prod_m \Delta g_m$; hence the entropies vary as the logarithm of this quantity, and $\tau(\xi, \eta; \varepsilon)$ grows as the logarithm of the uncertainty.

There remains the dependence on ε. For many K-mixing systems it is known that

$$H(\xi) - H(\xi \mid \eta_{-\infty}^{-t}) \sim e^{-kt} \qquad \text{as} \quad t \to \infty$$

Hence the dependence on ε is also logarithmic. This is true not only for Bernoulli and Markov dynamical systems, but also, e.g., for the continuous fractions transformation.[3] In fact, I do not know of any K-mixing system for which the convergence is not exponential, but I have tried in vain to prove that it is exponential in general. Prof. Sinai has told me that exponential convergence of entropy does not hold for a general K-mixing system; he has not given an explicit counterexample, but he suggested where some may be found. Therefore, I end with an open problem: find the class of K-mixing systems for which the convergence of entropies is exponential, and, in particular, find a counterexample to the exponential convergence.

REFERENCES

1. M. Reed and B. Simon, *Methods of Modern Mathematical Physics I* (Academic Press, New York, 1972), p. 62.
2. P. Billingsley, *Ergodic Theory and Information* (Wiley, New York, 1965).
3. P. Lévy, *Théorie de l'addition des variables aléatoires* (Paris, Gauthier-Villars, 1954), Chapter 9.