Kolmogorov entropy and numerical experiments

Giancarlo Benettin
Istituto di Fisica dell'Università, and Gruppo Nazionale di Struttura della Materia del Consiglio Nazionale delle Ricerche, Padova, Italy

Luigi Galgani
Istituto di Matematica and Istituto di Fisica dell'Università, Milano, Italy

Jean-Marie Strelcyn
Départment de Máthematiques Centre Scientifique et Polytechnique, Université Paris-Nord, Paris, France
(Received 8 June 1976)

Numerical investigations of dynamical systems allow one to give estimates of the rate of divergence of nearby trajectories, by means of a quantity which is usually assumed to be related to the Kolmogorov (or metric) entropy. In this paper it is shown first, on the basis of mathematical results of Oseledec and Piesin, how such a relation can be made precise. Then, as an example, a numerical study of the Kolmogorov entropy for the Hénon-Heiles model is reported.

I. INTRODUCTION

In recent years many attempts have been made in order to investigate the so-called stochasticity properties of dynamical systems, in particular Hamiltonian systems, by numerical computations. However, stochasticity is generally defined and tested in a rather qualitative way, and the connection between the empirical parameters introduced to describe it and rigorous theoretical concepts is far from being clear.

One of the most powerful empirical tools has always been the study of the divergence of nearby trajectories in phase space. Such a method allows one to define a quantitative parameter (the "entropylike quantity"), which is supposed to be strictly related to the Kolmogorov (or metric) entropy for associated flow.[1-4]

The aim of the present paper is to analyze this entropylike quantity, deriving its precise connection with the metric entropy, and to explain certain properties observed in the numerical computations. This connection turns out to be particularly simple for the case of Hamiltonian systems with two degrees of freedom. As an example, for one of them, the well-known Hénon-Heiles model,[5-7] we compute the entropylike quantity and test its properties; moreover, we are able to draw a tentative curve for the entropy itself as a function of energy.

In Sec. II we collect first the necessary mathematical tools, i.e., the results of Oseledec[8] and the fundamental results of Piesin.[9,10] [We are very grateful to Dr. A. B. Katok (Moscow) for the communication of the latter results.] We then recall the definition of the entropylike quantity and explain its empirically observed properties. The numerical example for the Hénon-Heiles model is treated in Sec. III.

This paper has been written, as far as possible, in a self-contained way; however, a certain familiarity with ergodic theory and in particular with entropy is necessary (see, for example, Refs. 11 and 12). The elementary notions on differentiable manifolds used here can be found, for example, in Ref. 13.

II. THEORETICAL ANALYSIS OF THE NUMERICAL COMPUTATIONS

A. Mathematical preliminaries: Lyapunov characteristic numbers and entropy

Let us give first the main definitions and fix the notation.

Let M be a differentiable, n-dimensional, compact, connected Riemannian manifold of class C^2. If $x \in M$, the tangent space to M at x and the norm induced in it by the Riemannian metric on M will be denoted by E_x and $\|\cdots\|$, respectively. Let X be a vector field of class C^2 defined on M and $\{T^t\}$ the flow induced by X, i.e., for any t let $T^t x = x(t)$, where $\{x(t)\}$ is an integral curve of the vector field X such that $x(0)=x$. The tangent mapping of E_x onto $E_{T^t x}$ induced by the diffeomorphism T^t will be denoted by dT^t_x. It will also be assumed that the flow $\{T^t\}$ preserves a normalized measure μ which is equivalent to the Lebesgue measure on M and whose density in local coordinates is of class C^2, i.e., that the flow $\{T^t\}$ admits an integral invariant of order n and class C^2.

The following theorems A and B, which partially summarize theorems 2 and 4 of Ref. 8, are the basis for all further considerations of the present paper.

Theorem A. There exists a measurable set $M_1 \subset M$, $\mu(M_1) = 1$, such that for every $x \in M_1$ the following properties hold:

(a) For every vector $e \in E_x$, $e \neq 0$, the limit

$$\lim_{t \to \infty} (1/t) \ln \|dT_x^t(e)\| = \lambda(x, e)$$

exists and is finite. M being compact, such a limit is independent of the Riemannian metric chosen on M.

(b) There exists a basis $(e_1, \ldots, e_n)$ of E_x such that

$$\sum_{i=1}^{n} \lambda(x, e_i) = \inf_{\Pi} \sum_{i=1}^{n} \lambda(x, \tilde{e}_i),$$

where

$$\Pi = \{(\tilde{e}_1, \ldots, \tilde{e}_n) : (\tilde{e}_1, \ldots, \tilde{e}_n) \text{ is a basis of } E_x\}.$$

As e varies in E_x, $\lambda(x, e)$ takes only the values $\{\lambda(x, e_i)\}_{1 \leq i \leq n}$.

The number $\lambda(x, e)$ is called the Lyapunov characteristic number of the vector $e \in E_x$, and the numbers $\lambda(x, e_i)$, which depend only on the flow $\{T^t\}$ and the point x, are called the Lyapunov characteristic numbers of the flow $\{T^t\}$ at x. We will use the notation $\lambda(x, e_i) = \lambda_i(x)$, $1 \leq i \leq n$, and suppose $\lambda_1(x) \leq \lambda_2(x) \leq \cdots \leq \lambda_n(x)$. Then the functions $\lambda_1, \ldots, \lambda_n$ are defined on M and can be shown to be measurable with respect to the Lebesgue measure on M.

We will also denote $\lambda_n(x)$ by $\lambda_{\max}(x)$. This is the quantity of main interest for the present paper, since it will be shown below to correspond to the entropylike quantity.

Given $x \in M_1$, the numbers $\{\lambda_i(x)\}_{1 \leq i \leq n}$ are not all necessarily distinct. Denote by $\{\nu_j(x)\}_{1 \leq j \leq s(x)}$ the distinct values taken by the numbers $\{\lambda_i(x)\}_{1 \leq i \leq n}$, and by $k_j(x)$ the multiplicity of $\nu_j(x)$. Also, let $\nu_i < \nu_j$ if $i < j$.

Theorem B. For every $x \in M_1$ there exist linear subspaces $H_1, \ldots, H_s$ of E_x, with $s = s(x)$, such that (a) $E_x = H_1 \oplus \cdots \oplus H_s$ ($\oplus$ denotes, as usual, direct sum); (b) $\dim H_j = k_j(x)$, $1 \leq j \leq s$; (c) if $e \neq 0$, $e \in H_j$, then

$$\lim_{t \to \pm\infty} (1/|t|) \ln \|dT_x^t(e)\| = \pm \nu_j(x), \quad 1 \leq j \leq s;$$

(d) if $e \neq 0$, $e \in H_1 \oplus \cdots \oplus H_j$ but $e \notin H_1 \oplus \cdots \oplus H_{j-1}$, then $\lambda(x, e) = \nu_j(x)$, $1 \leq j \leq s$.

From this theorem it immediately follows that if one chooses in E_x a vector e "at random" then one may expect to find $\lambda(x, e) = \nu_s(x) \equiv \lambda_{\max}(x)$. Indeed, the vectors e, such that $\lambda(x, e) < \nu_s(x)$, constitute a subspace of E_x which has positive codimension and thus vanishing Lebesgue measure.

We recall now the fundamental theorem of Piesin connecting Lyapunov characteristic numbers with metric entropy:

Theorem C. Under the conditions given above, one has

$$h_\mu\{T^t\} = \int_M \left[\sum_{\lambda_i(x) > 0} \lambda_i(x) \right] d\mu(x)$$

where $h_\mu(\{T^t\})$ is the entropy of the flow $\{T^t\}$ with respect to the invariant measure μ.

The following remark is an immediate consequence of the very definition of the Lyapunov characteristic numbers: Suppose the vector field X does not have singular points, i.e., for every $x \in M$ it is $X(x) \neq 0$. Then for any $x \in M_1$ one has $\lambda(x, X(x)) = 0$. As a consequence, at least one of the Lyapunov characteristic numbers of the flow $\{T^t\}$ vanishes for every $x \in M_1$, so that one has $\lambda_{\max}(x) \geq 0$ for $x \in M_1$.

From the above remark and Piesin's theorem one thus gets the inequality

$$\int_M \lambda_{\max}(x)\, d\mu(x) \leq h_\mu(\{T^t\}) \leq (n-1) \int_M \lambda_{\max}(x)\, d\mu(x). \tag{1}$$

Let us now consider the particular case of a Hamiltonian flow of N degrees of freedom. We assume we are given a function $H(q, p)$ of class C^2 defined on an open subset U of R^{2N}, with $N \geq 1$ and $q = (q_1, \ldots, q_N)$, $p = (p_1, \ldots, p_N)$, such that $\operatorname{grad} H(q, p) \neq 0$ for $(q, p) \in U$. The Hamiltonian flow $\{T^t\}$ on U is thus defined by the Hamilton equations of motion with Hamiltonian H.

The energy surfaces $\Omega_E = \{(q, p) \in U : H(q, p) = E\}$ are not necessarily compact. We will suppose that there exists an interval of energies E such that Ω_E contains a $(2N-1)$-dimensional submanifold Γ_E which is compact, connected, and $\{T^t\}$ invariant.

Denote by $\{T_E^t\}$ the restriction of the flow $\{T^t\}$ to Γ_E. It is well known that the Lebesgue measure on R^{2N} induces on Γ_E an invariant measure μ_L (strictly positive and of class C^2 in local coordinates) which can be supposed to be normalized by $\mu_L(\Gamma_E) = 1$. Under such conditions inequality (1) becomes

$$\int_{\Gamma_E} \lambda_{\max}(q, p)\, d\mu_L(q, p) \leq h_{\mu_L}(\{T_E^t\}) \leq 2(N-1) \int_{\Gamma_E} \lambda_{\max}(q, p)\, d\mu_L(q, p). \tag{2}$$

Such an inequality will be the basis for the discussion of the connections between the entropylike quantity and entropy. A more stringent inequality, which holds under rather particular assumptions and will be of interest in Sec. III B, is deduced in Sec. II D.

In closing this subsection, we may remark that the results of Oseledec and Piesin recalled here for a flow $\{T^t\}$ are easily formulated also in the case of a diffeomorphism T.

B. Entropylike quantity

Here, we give the description of the quantities $k_n(\tau, x, d)$ and $k(\tau, x, d)$ which were defined in Ref. 4 and also in the cited works by Chirikov and coworkers.[1-3]

We make reference here to the case of a Hamiltonian flow $\{T_E^t\}$ on Γ_E, but the cases of a flow or a diffeomorphism on a smooth manifold would be treated in a similar way.

Given $\tau > 0$, consider the diffeomorphism T^τ of Γ_E onto itself. Fix a point $x \in \Gamma_E$ and another point $y \in \Gamma_E$ very close to x, not on the same trajectory. Denote by d the segment relaying x to y and by $|d|$ its length.

Let $x_1 = T^\tau x$ and $|d_1| = \|T^\tau x - T^\tau y\|$, where $\|\cdots\|$ is the Euclidean norm. Denote by y_1 the unique point of the half-line issuing from x_1 and containing $T^\tau y$ such that $\|y_1 - x_1\| = |d|$. Then one can iterate such a procedure and define $x_2 = T^\tau x_1 = T^{2\tau} x$, $|d_2| = \|T^\tau x_1 - T^\tau y_1\|$; by y_2 we will denote the unique point of the half-line issuing from $x_2 = T^\tau x_1$ and containing $T^\tau y_1$, such that $\|y_2 - x_2\| = |d|$, and so on (see Fig. 1).

One thus gets a sequence of positive numbers $\{|d_i|\}$, $i = 1, 2, \ldots$, and one can so define the quantity

$$k_n(\tau, x, d) = \frac{1}{n\tau} \sum_{i=1}^{n} \ln \frac{|d_i|}{|d|}.$$

From the numerical computations described in Ref. 4 it appears that for the model there considered, if $|d|$ is not too big, one finds (i) the limit $\lim_{n\to\infty} k_n(\tau, x, d) = k(\tau, x, d)$ seems to exist; (ii) $k(\tau, x, d)$ is independent of τ; and (iii) $k(\tau, x, d)$ is independent of d.

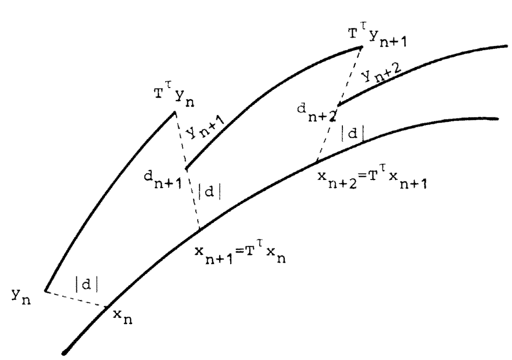

FIG. 1. Definition of the entropylike quantity.

In general, all of the numerical computations on a large class of Hamiltonian systems[1-4,7,14] show that for those systems, given an energy E, Γ_E decomposes, roughly speaking, into two regions which are $\{T_E^t\}$ invariant. One of those regions, which is called the ordered region (or sometimes the stable region), is characterized by the property that the trajectories of the flow $\{T_E^t\}$ have a behavior which seems similar to quasiperiodic motions. The other region, which is called the stochastic one, is instead characterized by a very irregular behavior of the trajectories of the flow, which seems similar to the behavior of the Anosov flows (see Ref. 1, and particularly Refs. 7 and 14). Such characterizations are not rigorous, but have nevertheless an undeniable heuristic value. It appears moreover that (iv) $k(\tau, x, d) = 0$ if x is taken in the ordered region of Γ_E; and (v) $k(\tau, x, d)$ is independent of the choice of x if x is taken in the stochastic region of Γ_E. In such a case, $k(\tau, x, d)$ is always positive. Property (v) allows one to speak simply of the quantity $k = k(E)$ instead of $k(\tau, x, d)$, with x belonging to the stochastic region of Γ_E. The number $k(E)$ was considered as an entropylike quantity. In Sec. II C such a statement will be given a precise meaning.

C. Identification of the entropylike quantity

We are now able to explain (i)–(iii) of Sec. II B and give some heuristic remarks about (iv) and (v). This will be based on the identification of $k(\tau, x, d)$ with the Lyapunov characteristic number $\lambda(x, e)$, where $e = y - x$.

(a) It is clear that, τ being fixed and $|d|$ sufficiently small, one has $|d_1|/|d| \cong \|dT_x^\tau(e)\|/\|e\|$, and moreover

$$\frac{|d_2|}{|d|} \cong \frac{1}{\|e\|} \left\| dT^\tau_{T^\tau_x}\left(\frac{\|e\|}{\|dT^\tau_x(e)\|} dT^\tau_x(e) \right) \right\| = \frac{\|dT^{2\tau}_x(e)\|}{\|dT^\tau_x(e)\|},$$

where the property $dT_x^{t+s} = dT^t_{T^s_x} dT^s_x$, following from $T^{t+s} = T^t T^s$, has been used. In general one has

$$\frac{|d_i|}{|d|} = \frac{\|dT_x^{i\tau}(e)\|}{\|dT_x^{(i-1)\tau}(e)\|}, \quad i = 1, 2, 3, \ldots.$$

As a consequence,

$$k_n(\tau, x, d) = \frac{1}{n\tau} \sum_{i=1}^{n} \ln \frac{|di|}{|d|}$$

$$\cong \frac{1}{n\tau} \sum_{i=1}^{n} \ln \frac{\|dT_x^{i\tau}(e)\|}{\|dT_x^{(i-1)\tau}(e)\|}$$

$$= \frac{1}{n\tau} \ln \frac{\|dT_x^{n\tau}(e)\|}{\|(e)\|},$$

and this allows, by note (a) of Theorem A, to identify $k(\tau, x, d) = \lim_{n\to\infty} k_n(\tau, x, d)$ with the Lyapunov characteristic number $\lambda(x, e)$. This identification is made with an error which tends to zero with $|d|$.

(b) Property (ii) of Sec. II B is then an immediate consequence of the very definition of the Lyapunov characteristic numbers.

(c) In (a) above $k(\tau, x, d)$ has been identified with $\lambda(x, e)$. Then, by the remark following Theorem B, one sees that if one chooses d at random, one may expect to find $k(\tau, x, d) = \lambda_{\max}(x)$.

(d) While properties (i)–(iii) of Sec. II B are consequences of the Oseledec theorem on the Lyapunov characteristic numbers and of the very definition of $k(\tau, x, d)$, properties (iv) and (v) instead have to be considered as empirical, and are far from being well understood. The following considerations are then, by necessity, of a heuristic character:

For the ordered region, the vanishing of the Lyapunov numbers should be related to the fact that its numerically computed trajectories are apparently of quasiperiodic type; this much we can say about (iv).

Property (v) supports instead the idea that the stochastic region either is ergodic or contains an ergodic component the measure of which is close to the measure of the stochastic region itself. In this connection we recall that an example of a flow, preserving Lebesgue measure, has been found[15] which is not ergodic, while all of its Lyapunov characteristic numbers but one are nonvanishing almost everywhere.

D. Connection with entropy

For the connection between entropy and the entropylike quantity, the only rigorous relation that can be given is essentially inequality (2), which follows from Piesin's theorem (Theorem C here) and the assumption that the vector field considered does not have singular points.

However, use may be made of the heuristic picture described in Secs. II B and II C concerning those particular models for which stochastic and ordered regions appear to exist. Indeed one may then assume that $\lambda_{\max}(x)$ vanishes in the ordered region of Γ_E and is equal to a positive constant, $k(E)$, in the stochastic region S_E of Γ_E, so that inequality (2) gives

$$\mu_L(S_E)k(E) \le h_{\mu_L}(\{T_E^t\}) \le 2(N-1)\mu_L(S_E)k(E) . \quad (3)$$

A more stringent relation, to be used in Sec. III B, 7, can, however, be given for the Hamiltonian systems such that $H(q,p) = H(q, -p)$, a condition which is satisfied in all models to which reference has been made in this paper. Indeed, defining $\psi: \Gamma_E \to \Gamma_E$ by $\psi(q,p) = (q, -p)$, if one assumes that the stochastic region S_E is ψ invariant and that the functions $\lambda_{N+1}, \ldots, \lambda_{2N-1}$ are μ_L constant almost everywhere on S_E, then one easily deduces, on S_E, the relations

$$\lambda_{2N-k} = -\lambda_k, \quad k = 1, \ldots, N \quad (4)$$

(in particular $\lambda_N = 0$). This is seen as follows: Let $x = (q,p)$; from $H(\psi(x)) = H(x)$, as is well known, one gets the time-reversal property $T_E^{-t} = \psi T_E^t \psi^{-1}$ of the flow. Then from part (c) of Theorem B one gets

$$\lim_{t\to\infty} \frac{1}{|t|} \ln \|dT_x^t(e)\| = -\lim_{t\to-\infty} \frac{1}{|t|} \ln \|dT_x^t(e)\|$$

for all $0 \ne e \in H_j$, $1 \le j \le s$, so that one immediately deduces $\lambda(x, e) = -\lambda(\psi(x), d\psi(e))$; the stated property then follows.

As a consequence, inequality (3) can then be strengthened and one gets

$$\mu_L(S_E)k(E) \le h_{\mu_L}(\{T_E^t\}) \le (N-1)\mu_L(S_E)k(E) . \quad (5)$$

This is of particular interest in the case $N = 2$, because it then gives the approximate equality

$$h_{\mu_L}(\{T_E^t\}) \cong \mu_L(S_E)k(E) . \quad (6)$$

This relation will allow us to produce, for the Hénon-Heiles model, an approximate curve for entropy as a function of energy, on the basis of numerical estimates for $k(E)$ and $\mu_L(S_E)$.

For this model it was possible to check numerically the two additional hypotheses introduced to obtain relation (4).

III. NUMERICAL EXAMPLE: THE HÉNON-HEILES MODEL

A. Hénon-Heiles model

Here, we briefly recall the definition and describe the main properties of the Hénon-Heiles model.[5] This is a dynamical system of two degrees of freedom characterized by the Hamiltonian

$$H(q_1, q_2, p_1, p_2) = \tfrac{1}{2}(p_1^2 + p_2^2 + q_1^2 + q_2^2) + q_1^2 q_2 - \tfrac{1}{3} q_2^3 ,$$

where $q_1, q_2, p_1, p_2 \in R$; a flow on R^4 is thus defined by the Hamilton equations

$$\frac{dq_1}{dt} = p_1, \quad \frac{dq_2}{dt} = p_2,$$

$$\frac{dp_1}{dt} = -q_1 - 2q_1q_2, \quad \frac{dp_2}{dt} = -q_2 + q_2^2 - q_1^2 .$$

It can be easily shown that the energy surfaces Ω_E are such that for $0 < E < \frac{1}{6}$ they admit a unique nonvoid invariant compact three-dimensional mani-

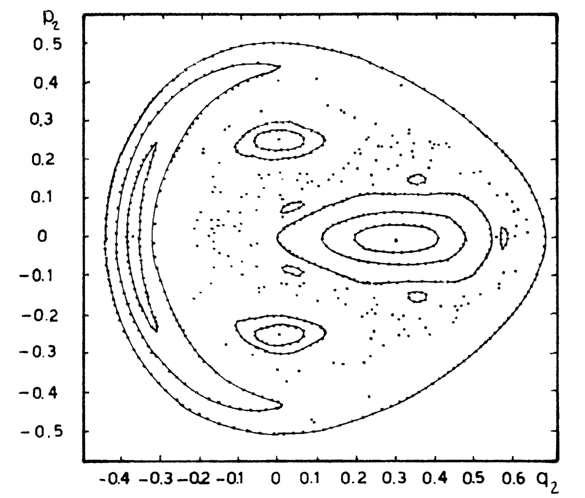

FIG. 2. Surface of section for the Hénon-Heiles model, at the intermediate energy $E = 0.125$, where ordered and stochastic regions have comparable measure (from Ref. 5).

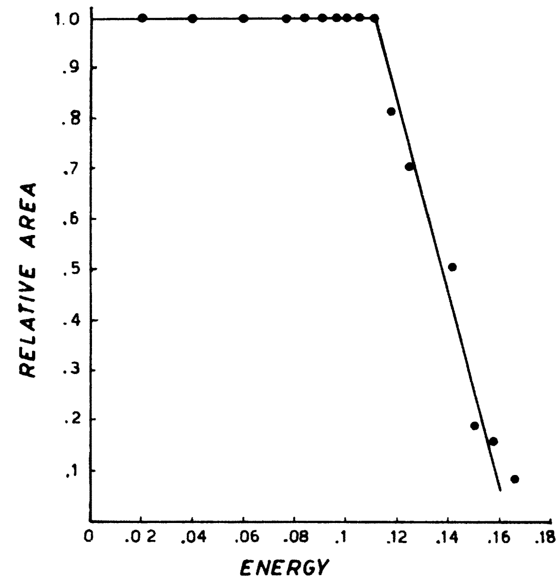

FIG. 3. Graph of the function $1-\tilde{\mu}(E)$ (from Ref. 5).

fold Γ_E and that the vector field X_H there defined does not have singular points.

This was the first model for which numerical computations indicated the existence of ordered and stochastic regions on Γ_E, according to the general qualitative description mentioned in Sec. II B.

This picture is clearly illustrated by Fig. 2, which is reproduced from Ref. 5. Such a figure gives a graphical presentation for the numerical integration of the equations of motion, by a standard device which reduces the study of a three-dimensional flow on Γ_E to the study of a two-dimensional plane mapping, as it will be now recalled.

Considering in Γ_E the two-dimensional surface given by $q_1 = 0$, one plots the successive points at which a particular solution intersects this surface with $p_1 > 0$. If one eliminates p_1 with the help of $H(q,p) = E$ and sets $q_1 = 0$, one can use q_2 and p_2 as coordinates on this two-dimensional surface; since $p_1^2 \geq 0$, they are restricted to belong to the region

$$\tilde{\Gamma}_E = \{(q_2, p_2): \tfrac{1}{2}(p_2^2 + q_2^2) - \tfrac{1}{3}q_2^3 \leq E\}.$$

As is seen from Fig. 2, which refers to $E = 0.125$, it appears that for some initial conditions these successive points are organized on various closed curves in the plane q_2, p_2. Other initial conditions (in the stochastic region) give instead successive points which are scattered around without any apparent order; actually, all such points in the figure refer to a single particular solution.

Let $\tilde{\mu}_L$ denote the (two-dimensional) Lebesgue measure on $\tilde{\Gamma}_E$ normalized by $\tilde{\mu}_L(\tilde{\Gamma}_E) = 1$. Let $\tilde{\mu}(E) = \tilde{\mu}_L(\tilde{S}_E)$, where $\tilde{S}_E$ is the stochastic region in $\tilde{\Gamma}_E$. Hénon and Heiles gave a numerical estimate for $\tilde{\mu}(E)$ as a function of E and found for $1 - \tilde{\mu}(E)$ the graph of Fig. 3. In this connection one may recall that as has been recently proven[16] one has rigorously $\tilde{\mu}(E) < 1$ for $0 < E < \frac{1}{6}$.

B. Results of numerical computations

We come now to the description of our numerical results. The computations were performed on a CDC 7600 computer, with a precision of 14 digits. The integration algorithm was the so-called central-point method,[17] correct up to the third order of the time step; the time step was typically 0.004.

The aim was to compute the quantity $k_n(\tau, x, d)$ described in Sec. II B. Given a value of E, the initial point x was chosen in $\tilde{\Gamma}_E$, i.e., by arbitrarily fixing q_2 and p_2, taking $q_1 = 0$, and determining p_1 by the condition $H(q,p) = E$. The displaced point y was chosen at a distance $|d|$ from x, with typically $|d| = 3 \times 10^{-4}$. Typically it was $\tau = 0.2$ and n up to 10^5.

The properties (i)–(v) of k_n discussed in Sec. II B and II C were checked with good accuracy.

For the independence from τ [property (ii) of Sec. II B], we remark that for any n the property

$$k_{jn}(\tau/j, x, d) \cong k_n(\tau, x, d), \quad j = 2, 3, \ldots,$$

should be expected to hold, from the considerations of Sec. II C. For example, this property was satisfied with an error of about 5×10^{-5} for $j = 2$ and 5×10^{-4} for $j = 10$, for any n up to 10^5, in a computation with x_0 in the stochastic region and $E = 0.125$. These values correspond to percentage errors of about 0.1% and 1%, respectively.

Analogously, the property [see property (iii) of Sec. II B]

$$k_n(\tau, x, \alpha d) \cong k_n(\tau, x, d), \quad \alpha \neq 0,$$

is expected to hold for any n. As a typical example, for $\alpha = 2.2$ the error was of about 5×10^{-5}, and the percentage error was 0.1% for the initial conditions given above.

The independence from the direction of the displacement d is, instead, expected to hold only in the limit $n\to\infty$. Several checks have been made: In the above conditions, changing the direction of d we found a difference in k_n which decreased from about 20% for $n = 250$ to 2% for $n = 2500$ and to 0.2% for $n = 25\,000$ or greater.

We pass now to the dependence on n and x. At low enough energies, where according to Hénon and Heiles (see Fig. 3) the measure of the stochastic region is negligible, we always found that k_n decreased with n (at least for large enough n), as illustrated in Fig. 4, which refers to various initial conditions at $E = 0.08$. With large enough n all curves seem to approach a straight line in the log-log scale, which corresponds to a behavior of the type $k_n \cong \alpha n^{-\beta}$ ($\alpha, \beta > 0$). This very regular behavior strongly supports the conjecture that $\lim_{n\to\infty} k_n = 0$. In particular it seems to be a rather general rule that in this model one has $\beta \cong 1$, as can be seen by comparison with the dotted line in Fig. 4.

This general behavior for k_n does not change by increasing the energy, provided one chooses appropriately the initial point x in such a way that it is in the ordered region. This is illustrated in Fig. 5.

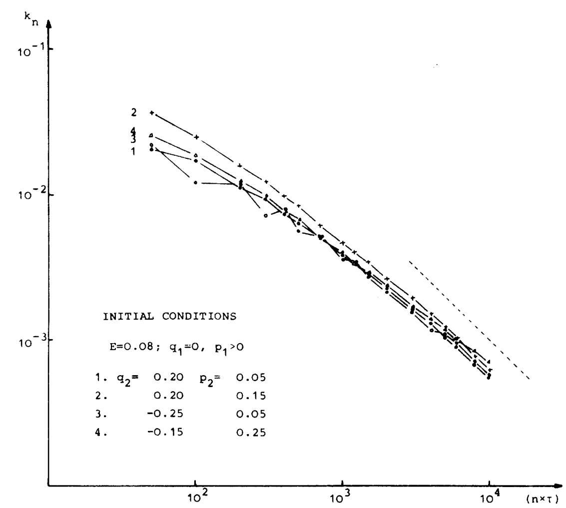

FIG. 4. Behavior of k_n at the fixed low energy $E = 0.08$, for different initial conditions.

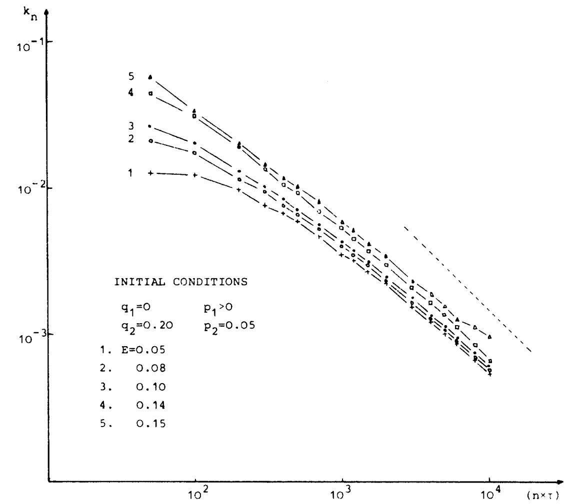

FIG. 5. Behavior of k_n at different energies, when the initial point is taken in the ordered region.

Let us now examine the case of a typical energy, $E = 0.125$, where the ordered and the stochastic regions have comparable measures (see Fig. 3). In Fig. 6 six curves are reported for $E = 0.125$. Curve 1, which has the same general feature of Figs. 4 and 5, corresponds to an initial point x lying in the large ordered region around the q_2 axis with $q_2 > 0$. Curve 2 refers to an initial point in one of the small islands surrounding such a region, while curve 3 corresponds to an initial point in one of the two symmetric ordered regions near the p_2 axis. For these curves, too, the limit seems to be zero, being approached, however, with a less regular behavior. Such a feature ap-

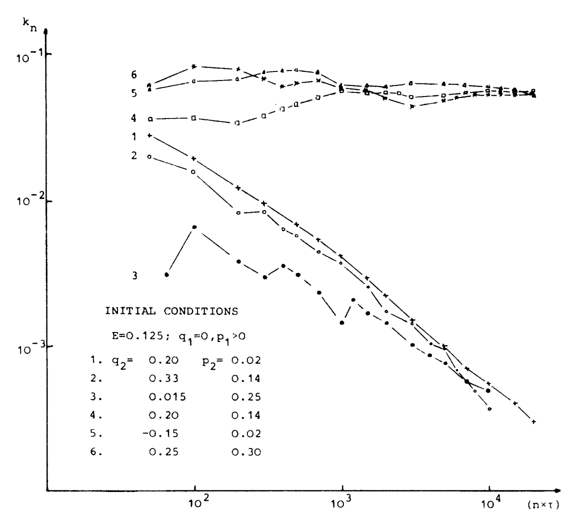

FIG. 6. Behavior of k_n at the intermediate energy $E = 0.125$, for initial points taken in the ordered (curves 1–3) or stochastic (curves 4–6) regions.

pears to be characteristic of these components of the ordered region, and is difficult to explain it rigorously at the moment. We hope to return to this problem in future. Curves 4–6 refer to initial points x in the stochastic region and rather clearly appear to approach a limit which is independent of x. This property has been checked for many other initial points at the same energy.

In general, for any energy $E>0$ we always found for the quantity $k=\lim_{n\to\infty} k_n$ either zero or a positive value which depends only on E. Such a positive value will be denoted, as in Sec. II B, by $k(E)$. The values found for $k(E)$ are reported in Fig. 7 by asterisks; the vanishing values found in the ordered region have also been marked (dots). All of the positive values between $E=0.105$ and 0.1666 are rather well defined with the same accuracy as that of Fig. 6. In particular, it may be remarked that two of them (i.e., those for $E=0.105$ and 0.110) refer to energies for which, according to Hénon and Heiles, the measure of the stochastic region is negligible (see Fig. 3). For the two points at $E=0.975$ and 0.1, they are defined with a much smaller accuracy, and a very careful search was necessary in order to find initial conditions with nonvanishing k at such energies. One can also note that apart from these two points the positive values are rather well fitted by an exponential, $k(E)=3.4e^{22E}$, as shown in Fig. 7 (continuous line).

We come now to an estimate for the metric entropy $h(E)=h_{\mu_L}(\{T_E^t\})$ as a function of energy E, on the basis of relation (6). The entropylike quantity $k(E)$ being known from the numerical computations described above, one needs only a knowledge of $\mu_L(S_E)$ as a function of E. For want of a direct estimate, we assume it to be given approximately by the function $\tilde{\mu}(E)=\tilde{\mu}_L(\tilde{S}_E)$ defined in Sec. IIIA and estimated numerically by Hénon and Heiles (Fig. 3). One thus has, approximately,

$$h(E)\cong\tilde{\mu}(E)k(E)\,. \qquad (7)$$

For example, if as a rough interpolation one takes for $\tilde{\mu}(E)$ the function given by $\tilde{\mu}(E)=0$ for $E<0.11$, $\tilde{\mu}(E)=1-17.6(E-0.11)$ for $0.11\le E\le\frac{1}{6}$, and for $k(E)$ the exponential $k(E)=3.4e^{22E}$, one gets the function $h(E)=0$ for $0<E<0.11$, $h(E)=60e^{22E}(E-0.11)$ for $0.11\le E\le\frac{1}{6}$, represented by a dotted line in Fig. 7. One may note that very probably one should have $h(E)>0$ for $0<E<\frac{1}{6}$, although this is still unproven, to our knowledge.

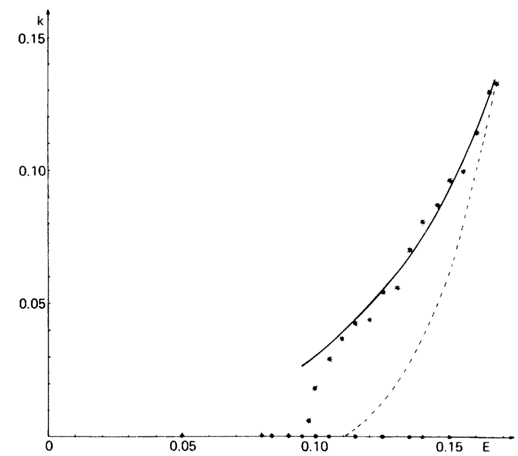

FIG. 7. Nonvanishing values found for k_n at different energies in the stochastic region (asterisks), vanishing values found in the ordered region (dots), exponential curve interpolating the nonvanishing values (continuous line), and tentative curve for entropy as a function of energy (dotted line).

IV. CONCLUDING REMARKS

We conclude by mentioning some open problems, strictly related to those considered in the present paper: (i) How can one calculate numerically the Lyapunov characteristic numbers other than the maximal one? (ii) Does the stochastic region in the Hénon-Heiles model contain more than one ergodic components, as suggested to us by Dr. M. Hénon? (iii) Does one have for the function $\tilde{\mu}(E)$ in the Hénon-Heiles model $\tilde{\mu}(E)>0$ for $0<E<\frac{1}{6}$, from both a numerical and a theoretical point of view? This last question is related to the already mentioned problem of whether one has $h(E)>0$ for $0<E<\frac{1}{6}$.

Note added in proof. Theorems A and B have also been proved by V. M. Millionščikov; see Mat. Sbornik 78, 179 (1969) [Math. USSR Sbornik 7, 171 (1969)].

[1]B. V. Chirikov: *Researches concerning the theory of nonlinear resonance and stochasticity.* CERN Transcript 71-40 Geneva, 1971 (unpublished).
[2]B. V. Chirikov and F. M. Izrailev, in Colloque International CNRS sur les Transformations Ponctuelles et leurs Applications, 10–14 Sept. (Toulouse, France, 1973) (unpublished).
[3]B. V. Chirikov, F. M. Izrailev, and V. A. Tayurski, Comput. Physics Commun. 5, 11 (1973).
[4]M. Casartelli, E. Diana, L. Galgani, and A. Scotti, Phys. Rev. A 13, 1921 (1976).
[5]M. Hénon and C. Heiles, Astron. J. 69, 73 (1964).

[6]Jurgen Moser and Walter T. Kyner, *Lectures on Hamiltonian Systems and Rigorous and Formal Stability of Orbits about an Oblate Planet* (Amer. Math. Soc., Providence, R. I., 1968).

[7]J. Ford, in *Fundamental Problems in Statistical Mechanics*, edited by E. D. G. Cohen (North-Holland, Amsterdam, 1975), Vol. 3.

[8]V. I. Oseledec, Tr. Mosk. Mat. Obsch. 19, 179 (1968) [Trans. Mosc. Math. Soc. 19, 197 (1968)].

[9]Ya. B. Piesin, Dokl. Akad. Nauk. SSSR 226, 774 (1976).

[10]Ya. B. Piesin, (to be published).

[11]P. Billingsley, *Ergodic Theory and Information* (Wiley, New York, 1965).

[12]P. Walters, *Ergodic Theory–Introductory Lectures: Lecture Notes in Mathematics No. 458* (Springer, Berlin, 1975).

[13]S. Sternberg, *Lectures on Differential Geometry* (Prentice Hall, Englewood Cliffs, 1964).

[14]J. Ford, *Stochastic Behavior in Nonlinear Oscillator Systems: Lectures in Statistical Physics No. 28* (Springer, Berlin, 1974).

[15]Ya. B. Piesin, Funkt. Anal. Jego Prilog 8, 81 (1974) [English translation in Funct. Anal. Applic. 8, 263 (1974)].

[16]M. Braun. J. Diff. Eq. 13, 300 (1973).

[17]See, for example, L. Verlet, Phys. Rev. 159, 98 (1967).