Geometrical hints for a nonperturbative approach to Hamiltonian dynamics

Marco Pettini*
Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125 Firenze, Italy
(Received 30 June 1992)

In this paper it is shown that elementary tools of Riemannian differential geometry can be successfully used to explain the origin of Hamiltonian chaos beyond the usual picture of homoclinic intersections. This approach stems out of first principles of mechanics and fundamental tools of Riemannian geometry. Natural motions of Hamiltonian systems can be viewed as geodesics of the configuration-space manifold M equipped with a suitable metric g_J, and the stability properties of such geodesics can be investigated by means of the Jacobi–Levi-Civita equation for geodesic spread. The study of the relationship between chaos and the curvature properties of the configuration-space manifold is the main concern of the present paper and is carried out by numerical simulations. Two different mechanisms for chaotic instability are found: (i) the trajectories are "scattered" by random encounters of regions of negative curvature (either scalar or Ricci curvature—it depends on the averaging procedure adopted); (ii) the "bumpiness" of (M,g_J) yields oscillations of the Ricci curvature along the geodesics so that parametric resonance makes them unstable also in regions of *positive* curvature. The geometric approach is intrinsically *nonperturbative* because everything is well defined at any energy, i.e., quasi-integrability is not required as in the case of classical perturbation theory. Therefore this approach is fit to describe the existence of the strong-stochasticity threshold (SST) in high-dimensional Hamiltonian flows. This threshold refers to a transition between weak and strong chaoticity of the dynamics and, correspondingly, between slow and fast mixing in phase space. In view of applications to equilibrium and nonequilibrium statistical mechanics, the SST appears as the transition feature of high dimensional Hamiltonian flows with the greatest physical significance. As the SST concerns chaotic dynamical behaviors, its existence cannot be understood within the framework of classical perturbation theory, whereas it is shown that the SST can be related to some major geometrical change of the constant-energy surfaces of phase space. A clear-cut distinction can be made between integrable and nonintegrable systems. Finally, a new meaning is given to the standard algorithm to compute numerical Lyapunov exponents, and it is shown that Oseledets multiplicative theorem is not necessary to justify it.

PACS number(s): 05.45.+b, 02.40.−k, 05.20.−y

I. INTRODUCTION

The subject of the present paper is nonlinear Hamiltonian dynamics in high-dimensional phase space. The first historical motivation for the study of these systems stems from celestial mechanics as it is witnessed by Poincaré's *Méthodes Nouvelles de la Méchanique Celeste* [1].

Motions of celestial bodies are regular, at least on our observational time scales. Planetary orbits in our solar system are mutually perturbed but regular. Therefore, the analytical techniques developed to describe these systems assume from the beginning the physical evidence of dynamical regularity. This is the physics one recognizes behind the existing analytical tools to tackle nonlinear Hamiltonian dynamics. The ensemble of these methods is known as classical perturbation theory (CPT).

In the *Méthodes Nouvelles* one can find also a description of the structure of the orbits in the vicinity of hyperbolic points where the perturbative series are divergent and where the so-called homoclinic intersections take place [2]. This is the origin of Hamiltonian chaos. On the other hand, CPT fails in quantitatively describing these chaotic regions of phase space just because of the divergence of perturbative series. (It could be argued that, for instance, Melnikov's method [3] is of perturbative character; anyway, at large perturbations or at high dimensionality, it is not very useful.)

Since the beginning of the 1950s, CPT has undergone a renewal of interest and a remarkable conceptual development. This can be attributed to the birth of the Kolmogorov-Arnold-Moser (KAM) theory which actually was a breakthrough.

CPT and thus KAM theory, too, deals with quasi-integrable systems, i.e., those described by Hamiltonians such as

$$H(\mathbf{p},\mathbf{q})=H_0(\mathbf{p})+H_1(\mathbf{p},\mathbf{q}),\quad \mu=\frac{\|H_1\|}{\|H_0\|}\ll 1\ . \tag{1}$$

Our experimental knowledge suggests that quasi-integrability is a good approximation for the dynamical behavior of planetary systems, satellites and so on, or for the structure of magnetic surfaces in tokamaks (in absence of disruptions) or for the beam behavior of particle accelerators, just to give some examples.

Conversely, ergodicity and mixing are generic empirical properties of the microscopic dynamics of condensed matter (solids and fluids). For this reason a statistical description is justified and effective.

Therefore celestial mechanics and statistical mechanics

describe opposite dynamical situations. Moreover, while the former usually deals with *dilute* few-body systems, the latter describes *dense* N-body systems with N of the order of the Avogadro number, so one would expect that such a big difference must be reflected by their respective descriptions of the dynamics.

A well-known theorem by Poincaré [1] and by Fermi [4] excluded, under some general conditions, the existence of smooth integral invariants besides the energy at $N \geq 3$. This implies that for generic nonintegrable systems, as in Eq. (1), the whole constant-energy surface in phase space is topologically accessible.

This theorem seems to be responsible for the common belief that the ergodic hypothesis at the basis of classical statistical mechanics is correct [5]. Doubts about this point of view have been raised independently by the famous numerical experiment by Fermi, Pasta, and Ulam (FPU) [6], and by the KAM theorem [7]. In the FPU experiment no tendency toward equipartition of energy was observed where it was expected, that is, in a chain of linear oscillators coupled by a weak anharmonicity.

On the other hand, the KAM theorem pointed out that a positive measure of invariant tori can survive even in the presence of a nonintegrable perturbation, provided that $\mu < \mu_c$, with μ_c some critical amplitude. A lot of work followed; it can be schematically summarized as follows.

From the analytical standpoint, it is now clear that the estimates of the critical perturbation amplitude μ_c give exceedingly small values; moreover, μ_c has a strong (at least exponential [8]) dependence upon the number N of degrees of freedom. Typically $\mu_c \sim \exp(-N \ln N)$. So it is doubtful whether the KAM theorem can have any physical relevance even at few degrees of freedom, and almost certainly it does not represent a problem for the foundations of statistical mechanics. A remarkable improvement of KAM theory started with Nekhoroshev theorem [9]. This one deals with *finite* time stability of regular orbits in phase space instead of *infinite* time stability as in the KAM theory. Again the perturbation amplitude μ must be smaller than some critical value μ_c that turns out very small, and the estimate of the lower bound for the stability time $\tau = \tau_0 \mu^{-1} \exp[(1/\mu)^\gamma]$ drops down to ~ 1 already at small N because $\gamma(N) \sim 1/N$ [10].

From the numerical standpoint, after a pioneering work on the dynamics of a chain of atoms coupled by a Lennard-Jones potential [11], we know that some kind of stochasticity threshold exists also for high-dimensional Hamiltonian flows. Below this threshold (some critical energy per degree of freedom) the motion is apparently regular, while above the threshold the motion appears chaotic and mixing. Moreover, this threshold seems independent (or at least weakly dependent) on N. Also in the case of the FPU model something similar happens [12] and this could seem to be an explanation of the result by Fermi and collaborators. Such an explanation was already proposed in Ref. [13], where the results of the FPU experiment were related with Kolmogorov's theorem. In fact, the lack of equipartition in the FPU experiment was attributed to the nonfulfillment of some stochasticity condition by the parameters chosen.

However, recent extensive numerical simulations [14] have clearly shown that—within the energy range explored—equipartition of energy is *always* attained in the FPU model, and that what was called the stochasticity threshold or equipartition threshold [12] in reality is a *strong-stochasticity threshold* (SST). On the other hand, at $N \geq 3$ there is no topological obstruction to the complete filling of phase space (Poincaré-Fermi theorem), and nonergodicity of metrical origin is only possible below the KAM threshold. As a consequence, physically relevant phenomena are expected essentially from the study of *relaxation times* or *mixing rates.*

At energies larger than the SST, fast relaxation and fast mixing are observed. On the contrary, at energies lower than the SST, very long relaxation times can be obtained by reducing the energy of the initial excitation; anyway, the largest Lyapunov exponent λ_1 is always found positive, which means that some chaoticity is always present.

A crossover in the scaling behavior of $\lambda_1(\varepsilon)$—where ε is the energy per degree of freedom—provides the operational definition of the SST. The strongly chaotic regime corresponds to $\lambda_1(\varepsilon) \sim \varepsilon^{2/3}$. This can be explained with a random matrix approximation for the tangent dynamics [14]. At variance, the weakly chaotic regime corresponds to a steeper law: $\lambda_1(\varepsilon) \sim \varepsilon^2$.

Measuring small λ_1 requires very long integration times [14]. Sometimes erroneous conclusions can be drawn by insufficiently long runs yielding $\lambda_1 \sim t^{-1+\alpha}$, $\alpha \ll 1$, during very long trappings of the trajectories (see Sec. III of this paper).

It is by no means surprising that λ_1 is always positive for a nonintegrable system. In fact, the intersection of the resonant manifolds of H_0 with the constant-energy surface originates a complex, connected, and dense network in phase space, which is transformed into a stochastic web by an arbitrary weak perturbation H_1. But even if $\lambda_1(\varepsilon) > 0$ at any energy ε, and even in the hypothesis that the overwhelming majority of the phase-space trajectories is chaotic, still the topology of phase-space paths can change a lot by varying the energy. In fact the SST is ascribed to a transition from *slow* diffusion in phase space (in the language of resonances this means that diffusion mainly occurs *along* resonances), to *fast* diffusion *across* resonances.

This kind of transition is *global* (it does not depend on the initial condition) and so is *intrinsically nonperturbative,* i.e., it is not explained by the existing results of CPT, as already discussed in Ref. [14]. In fact, quasi-integrability—assumed from the very beginning in CPT—is not a reasonable global property of systems whose phase space is mostly chaotic (of course a local perturbative description is always allowed on a time scale much shorter than the dynamical instability time).

It is worth noticing that a numerical work [15] on the classical dynamics of the Heisenberg planar XY model suggests that existence of an intriguing relationship between the SST and the dynamical counterpart of phase transitions. In fact a SST is found at the critical temperature T_c, at which the system undergoes a Kosterlitz-

Thouless phase transition; here the threshold is given by a crossover in the scaling of $\lambda_1(T)$.

This is to say that the dynamical properties of high-dimensional Hamiltonian flows have possibly a wider interest to equilibrium statistical mechanics than the *ergodic problem* at its foundations.

In the present paper we investigate the possibility of describing the chaotic component of phase space by means of elementary tools of Riemannian differential geometry. The aim is mainly twofold: (i) to suggest an explanation of the existence of the SST, and (ii) to provide, at least in principle, a statistical-mechanical algorithm to compute this threshold.

The idea of looking at Riemannian differential geometry as an analytical alternative to CPT comes from the observation that the only theoretical framework dealing with the opposite situation to quasi-integrability is *ergodic theory*. An important contribution to this topic was given by Krylov. He realized how tightly the mixing processes of statistical systems are related to the exponential instability of the underlying dynamics [16]. Moreover, Krylov knew the works by Hadamard [17], Hedlund [18], and Hopf [19] on the stability properties of the geodesics of compact Riemannian manifolds with constant negative curvature, and he was the first physicist who grasped the relevance of these mathematical works for the understanding of phase-space mixing.

The follow-up of Krylov's ideas took place within the ergodic theory, with the fundamental contributions by Anosov, Sinai, and others on geodesic flows [20], whereas generic Hamiltonian flows, for instance like the FPU model, have not been touched by these methods with only few exceptions [21]. In what follows we show how Krylov's original intuitions can be developed with the aid of computer simulations shining a new light on the origin of weak and strong chaos—thus of slow and fast mixing—in high-dimensional Hamiltonian flows.

Section II is devoted to a brief recall of the geometrical description of mechanical systems with Riemann and Finsler spaces and contains a result by Eisenhart that is generally unknown. This section also contains a discussion about some approximations that can be made on the Jacobi—Levi-Civita equation for the geodesic deviation in order to define geometric indicators of chaos. The relationship with Lyapunov exponents and with a widespread method for their numerical computation is also discussed. Section III is devoted to a critical discussion of the results of numerical simulations on the FPU model, on the classical lattice φ^4 model, and on the Toda lattice.

A discussion on the perspectives of future development of a differential geometrical approach to Hamiltonian chaos is given in Sec. IV, where some conclusions are also drawn.

II. GEOMETRY AND HAMILTONIAN DYNAMICS

We aim at considering Newtonian mechanics from a geometrical point of view. More precisely, the trajectories of a Hamiltonian flow are regarded as geodesics of a Riemannian manifold, equipped with a suitable metric, so that weak or strong chaotic instability of the trajectories can be related with geometrical properties of the underlying manifold.

Certainly the natural geometrical setting of Hamiltonian dynamics is within the framework of symplectic or, more generally, Poisson geometry. But since in physics we are mostly interested in the study of Hamiltonian systems that have a standard kinetic part, also Riemannian geometry can be used. The advantage of using Riemannian manifolds relies on the possibility of measuring the distance between two points, thus the separation between two different trajectories. It is well known how Newtonian mechanics can be rephrased in the Riemannian geometrical language; nevertheless, it is not out of place to recall the main concepts and definitions that are used to describe and measure Hamiltonian chaos from a differential geometrical point of view.

First of all let us mention that different choices are possible for the ambient space (configuration space, configuration space-time, phase space), therefore different metrics can be used. Besides, for more general problems, when the Riemannian description is not possible, still a geometrical approach in the same spirit can be developed with the aid of Finsler spaces.

In the following, using the configuration-space manifold equipped with a Riemannian metric, only a first step is done in the direction of a geometrical description of Hamiltonian chaos.

The relationship between Riemannian geometry and Newtonian mechanics has its origin in their variational formulations. In fact, the geodesics of a Riemannian manifold are defined as extremals of the arc-length functional

$$l_{AB} = \int_A^B ds \ , \tag{2}$$

and the trajectories of a mechanical system are given by the extremals of the action integral

$$\mathcal{A} = \int_{t_0}^{t_1} L(q^i, \dot{q}^i) dt \tag{3}$$

according to Hamilton's least action principle, or by the extremals of

$$A = \int_{\gamma(t)} 2T(q^i, \dot{q}^i) dt \tag{4}$$

according to Maupertuis's principle [T is the kinetic energy and $\gamma(t)$ are all the isoenergetic curves joining two given points $\mathbf{q}_0$ and $\mathbf{q}_1$]. Thus the trajectories of a mechanical system can be viewed as geodesics of a manifold, once a convenient metric is given to this manifold.

A. Maupertuis's principle and the kinetic-energy metric on M

Let us begin with the case where the metric is provided by kinetic energy. In this case we have at our disposal a positive-definite quadratic form which defines a proper Riemannian manifold [22].

Now, consider as ambient space a smooth n-dimensional manifold M, called configuration space, where the motion takes place. Let $TM = \cup_{q \in M} T_q M$ be

its tangent bundle. The function L: $TM \to \mathbb{R}$, defined by $L = \frac{1}{2}\langle v, v \rangle$, is a free Lagrangian describing a free motion on M. In this case $L \equiv T$, i.e., the Lagrangian has only a kinetic part. The scalar product $\langle v, v \rangle$ for all the tangent vectors v provides a Riemannian metric on M. In local coordinates one has $L = g_{ij}\dot{q}^i\dot{q}^j$, where g_{ij} is the metric tensor on M.

Let ω_L be the Lagrangian closed two-form on TM; ω_L is associated to the canonical symplectic form $\omega_0 = \sum_{i=1}^n dq^i \wedge dp_i$ defined on T^*M (phase space) by means of the Legendre transform $\mathbf{F}L$: $TM \to T^*M$, so $\omega_L = (\mathbf{F}L)^*\omega_0$. In local coordinates, putting $L_{\dot{q}^i\dot{q}^j} = (\partial^2 L / \partial\dot{q}^i \partial\dot{q}^j)$, one has $\omega_L = \sum_{i,j=1}^n (L_{\dot{q}^i q^j} dq^i \wedge dq^i + L_{\dot{q}^i\dot{q}^j} dq^i \wedge d\dot{q}^j)$. We denote by X_E the unique Lagrangian vector field on TM such that

$$\omega_L(X_E, Y) = dE(Y) \tag{5}$$

for each arbitrary vector field Y on TM, E is the "energy" given by $E = S - L$ with S: $TM \to \mathbb{R}$ defined by $S(v_x) = \mathbf{F}L(v_x)v_x$.

The solutions of the Euler-Lagrange equations are the natural motions of the system described by the Lagrangian L. Besides, the natural motions are the integral curves of the vector field X_E. Notice that second-order equations are possible on TM but not on T^*M.

The so-called base integral curves of X_E are given by the canonical projection of the integral curves of X_E from TM to M. Conversely, for each curve γ_0: $\mathbb{R} \to M$ a natural lift exists from M to TM, that is, $\gamma_0 \to (\gamma_0, \dot{\gamma}_0)$.

It can be proved that γ_0: $\mathbb{R} \to M$ is a base integral curve of X_E if and only if γ_0 is a geodesic for M, that is, $\nabla_{\dot{\gamma}_0}\dot{\gamma}_0 = 0$ with $\nabla_{\dot{\gamma}_0}$ the covariant derivative of the canonical Levi-Civita connection associated to g_{ij}.

In local coordinates, having set $\gamma_0(s) = (q^1(s), \ldots, q^n(s))$, one has

$$\frac{d^2q^i}{ds^2} + \Gamma^i_{jk}\frac{dq^j}{ds}\frac{dq^k}{ds} = 0 , \tag{6}$$

where s is the proper time and—as usual—the summation over repeated indices is implicitly assumed; Γ^i_{jk} are the Christoffel coefficients of the Levi-Civita connection associated with g_{ij} and are given by

$$\Gamma^i_{jk} = \frac{1}{2}g^{im}\left[\frac{\partial g_{mk}}{\partial q^j} + \frac{\partial g_{mj}}{\partial q^k} - \frac{\partial g_{jk}}{\partial q^m}\right] . \tag{7}$$

Now, if V: $M \to \mathbb{R}$ is a potential-energy function on M, we can incorporate it into the Lagrangian by defining

$$L_V = \frac{1}{2}\langle v, v \rangle - V(\pi_M v) , \tag{8}$$

where π_M: $TM \to M$ is the canonical projection of the tangent bundle, and then define energy as $E(v) = \frac{1}{2}\langle v, v \rangle + V(\pi_M v)$. Then $\gamma_0(s)$ is a base integral curve of the corresponding Lagrangian vector field X_E if and only if $\nabla_{\dot{\gamma}_0}\dot{\gamma}_0(s) = -\mathrm{grad}V[\gamma_0(s)]$, that is, in local coordinates,

$$\frac{d^2q^i}{ds^2} + \Gamma^i_{jk}\frac{dq^j}{ds}\frac{dq^k}{ds} = -g^{ij}\frac{\partial V}{\partial q^j} , \tag{9}$$

that are Euler-Lagrange equations.

Assuming M to be a compact manifold there exists a number E such that $E > V(\mathbf{q})$ for $\mathbf{q} \in M$. Then with such a number E one can associate a kinetic-energy metric or Jacobi metric on M by putting $\tilde{g} = [E - V(\mathbf{q})]g$; evidently $\tilde{g}$ is conformally equivalent to g, and in coordinates

$$\tilde{g}_{ij} = [E - V(\mathbf{q})]g_{ij} . \tag{10}$$

It can be shown that the base integral curves of the Lagrangian (8) coincide with geodesics of the Jacobi metric (10) up to a reparametrization with energy 1. In the following we shall denote the Jacobi metric by g_J.

If we denote by $\tilde{\Gamma}^i_{jk}$ the connection coefficients derived from the metric (10), the corresponding geodesics are given by

$$\frac{d^2q^i}{ds^2} + \tilde{\Gamma}^i_{jk}(\mathbf{q})\frac{dq^j}{ds}\frac{dq^k}{ds} = 0 . \tag{11}$$

Let us restrict to those systems whose kinetic-energy term, with a suitable choice of local coordinates, can be diagonalized, and so let us assume that $\tilde{g}_{ij} = [E - V(\mathbf{q})]\delta_{ij}$, hence, with $W = E - V$ and $W_i = \partial W / \partial q^i = -\partial V / \partial q^i$, Eq. (11) gives

$$\frac{d^2q^i}{ds^2} + \frac{1}{2W}\left[2W_j\frac{dq^j}{ds}\frac{dq^i}{ds} - \tilde{g}^{ij}W_j\tilde{g}_{kl}\frac{dq^k}{ds}\frac{dq^l}{ds}\right] = 0 ; \tag{12}$$

finally, using $ds^2 = 2W^2dt^2$, this yields

$$\frac{d^2q^i}{dt^2} = -\frac{\partial V}{\partial q^i} , \tag{13}$$

i.e., Newton's equations associated to L_V of Eq. (8).

B. Hamilton's principle and Finslerian metric on $M \times \mathbf{R}$

In order to get a regular variational problem from Hamilton's least action principle $\delta\mathcal{A} = 0$, the Lagrangian function $L(q^i, \dot{q}^i)$ has to fulfill some conditions. In particular, the variational problem has to be invariant for time reparametrization; that is, the integral curves must make stationary the action integral in Eq. (3), independent of the choice of the time parameter t. In order to fulfill this request, it is necessary and sufficient that the function $L(q^i, \dot{q}^i)$ be homogeneous of degree 1 in the velocities, that is, $L(q^i, \lambda\dot{q}^i) = \lambda L(q^i, \dot{q}^i)$, $\lambda > 0$. This condition is not so stringent as it could appear; in fact, by adding a supplementary dimension it is possible to define a parametrically invariant extension of the initial Lagrangian as follows [23]:

$$\Lambda(q^a, \dot{q}^a) = L(q^i, \dot{q}^i/\dot{q}^{n+1})\dot{q}^{n+1}, \quad a = 1, \ldots, n+1 , \tag{14}$$

where the parametric representation $q^a(u)$ is expressed as a function of an arbitrary parameter u. Hence $\dot{q}^a = dq^a/du$ and $dq^i/dt = \dot{q}^i/\dot{q}^{n+1}$. Equation (14) is homogeneous of degree 1 in the velocities; then taking u as an integration variable, the Hamiltonian action is

given by

$$\mathcal{S}=\int_{q_0}^{q_1}\Lambda(q^a,\dot{q}^a)du \ , \tag{15}$$

where the explicit expression of Λ reads as

$$\Lambda(q^a,\dot{q}^a)=\tfrac{1}{2}a_{ij}\frac{\dot{q}^i\dot{q}^j}{\dot{q}^{n+1}}-V\dot{q}^{n+1} \ . \tag{16}$$

In such a way the trajectories of our system in configuration space-time are given by the extremals of the functional (15), and the formalism is invariant with respect to reparametrizations.

The function $\Lambda(q^a,\dot{q}^a)$ naturally provides the tangent bundle $T(M\times\mathbb{R})$ with a metric function. The metric tensor $g_{\alpha\beta}(q^a,\dot{q}^a)$ defined through the metric function Λ as

$$g_{\alpha\beta}=\frac{1}{2}\frac{\partial^2\Lambda^2}{\partial\dot{q}^\alpha\partial\dot{q}^\beta} \tag{17}$$

provides the manifold $\mathcal{M}=M\times\mathbb{R}$ with a Finslerian structure; similarly to the case of a Riemannian manifold, from the knowledge of the metric tensor $g_{\alpha\beta}$ many important geometric properties of the manifold can be derived.

It is remarkable that the Finslerian geodesics on $\mathcal{M}$, associated to the parametrically extended Lagrangian Λ, are described by

$$\begin{aligned}&\frac{dp_i}{du}-\dot{q}^{n+1}\frac{\partial L}{\partial q^i}=0, \quad i=1,\ldots,n \ ,\\ &\frac{d}{du}\frac{\partial\Lambda}{\partial\dot{q}^{n+1}}=\frac{dH}{du}=0 \ ,\end{aligned} \tag{18}$$

where $p_a=\partial\Lambda/\partial\dot{q}^a$ and H is the Hamiltonian function associated with L. The first n equations, written in terms of the initial parameter t (i.e., $q^{n+1}=t$), are the equations of motion associated to the Lagrangian L, while the $(n+1)$th equation is nothing but the energy conservation along any geodesic.

Several definitions can be given [24] for the connection on a Finslerian manifold, though Cartan's connection has many advantages with respect to the others. Using Cartan's connection three different curvature tensors can be defined and also an equation for the vector field of geodesic variations can be introduced [24]; however, their derivation from the metric (17) is rather tedious when the metric function is given by the Lagrangian (16). Therefore, having mentioned the existence of an interesting direction for further investigations, we restrict ourselves to the simplest possibility in the Riemannian framework. Finsler spaces are necessary to describe Hamiltonian systems having a functional dependence on the momenta other than purely quadratic.

C. Eisenhart's metric on $M\times\mathbb{R}^2$

Consider as ambient space the configuration space-time, i.e., the $(n+1)$-dimensional product manifold $\mathcal{M}=M\times\mathbb{R}$. This space is the set of all the possible configurations of a mechanical system at different and continuously varying times. Time is a parameter in the Newtonian sense; thus $q^1,\ldots,q^n\in M$ and $q^0\in\mathbb{R}$, so that $q^0=t$.

Now, we could try to define a metric for $M\times\mathbb{R}$ by somehow relating ds to Ldt. Obviously this would give a pseudo-Riemannian metric.

In other words, the Lagrangian $L(q^a,\dot{q}^a)$, $a=0,1,\ldots,n$, can provide the quadratic form which is necessary to define a pseudo-Riemannian structure, in fact from $2L(q^a,\dot{q}^a)=a_{ij}\dot{q}^i\dot{q}^j-2V(q^a)(\dot{q}^0)^2$ one has

$$ds^2=a_{ij}dq^idq^j-2V(q^a)(dq^0)^3\equiv g_{\alpha\beta}(q^a)dq^\alpha dq^\beta \ . \tag{19}$$

Unfortunately the geodesic equations derived from (19) are

$$\ddot{q}^i+\left[\frac{\partial V}{\partial q_i}\right]-\frac{1}{2L}\frac{dL}{dt}\dot{q}^i=0 \ , \tag{20}$$

whose solutions do not coincide with natural motions because L is not a constant. Then a simple but nontrivial extension of the metric (19) was given by Eisenhart [25], who introduced an additional coordinate $q^{n+1}\in\mathbb{R}$ so that the time parametrization of the arc length becomes affine and the natural motions are obtained as geodesics of the new metric. Eisenhart's arc length is written as

$$ds_E^2=a_{ij}dq^idq^j-2V(q^a)(dq^0)^2+2dq^0dq^{n+1} \ . \tag{21}$$

Given two real constants k^2 and C, it is possible to compute explicitly [26] the function $q^{n+1}(t)$ as

$$q^{n+1}(t)=k^2t+C-\int_0^t dtL(q^a,\dot{q}^a) \ , \tag{22}$$

which shows how tightly q^{n+1} is related to the Hamiltonian action.

Eisenhart's theorem states that the trajectories—in configuration space-time $\mathcal{M}=M\times\mathbb{R}$—of a conservative, holonomic dynamical system, can be obtained as follows [26]: given two real constants k^2 and C, consider the product $\mathcal{M}\times\mathbb{R}$ equipped with Eisenhart's metric g_E defined by (21), where q^{n+1} is the coordinate along $\mathbb{R}$, then the dynamical trajectories are the projections on $\mathcal{M}$ of the geodesics of $(\mathcal{M}\times\mathbb{R},g_E)$ for which the arc length is positive definite and given by $ds_E^2=k^2dt^2$. Conversely, if any point $p\in\mathcal{M}$ of such a trajectory is associated to a point $p'\in\mathcal{M}\times\mathbb{R}$ whose projection coincides with p, and p' is defined by (22), then p' describes a geodesic of $\mathcal{M}\times\mathbb{R}$ fulfilling the preceding conditions.

Explicitly, g_E on $\mathcal{M}\times\mathbb{R}$ is given by

$$g_E=\begin{bmatrix}-2V(q^a) & \underline{0} & 1\\ \underline{0}^T & \underline{a} & \underline{0}^T\\ 1 & \underline{0} & 0\end{bmatrix} \tag{23}$$

and

$$g_E^{-1}=\begin{bmatrix}0 & \underline{0} & 1\\ \underline{0}^T & \underline{a}^{-1} & \underline{0}^T\\ 1 & \underline{0} & 2V(q^a)\end{bmatrix} \ , \tag{24}$$

where $\underline{0}$ is the null row vector of $\mathbb{R}^n$, $\underline{0}^T$ is its transpose, and $\underline{a}=\|a_{ij}\|$.

The nonvanishing connection coefficients associated with g_E can be easily computed, and for $a_{ij}=\delta_{ij}$ they are $\Gamma^i_{00}=\partial V/\partial q_i$ and $\Gamma^{n+1}_{0i}=-\partial V/\partial q^i$, so that the geodesic equations are

$$\frac{d^2q^i}{ds^2}+\Gamma^i_{00}(q^a)\frac{dq^0}{ds}\frac{dq^0}{ds}=0, \quad i=1,\ldots,n \ ,$$

$$\frac{d^2q^0}{ds^2}=0, \quad \frac{d^2q^{n+1}}{ds^2}+\Gamma^{n+1}_{0i}(q^a)\frac{dq^0}{ds}\frac{dq^i}{ds}=0 \ , \tag{25}$$

and since $ds^2=k^2dt^2$, one finds

$$\frac{d^2q^i}{dt^2}=-\frac{\partial V}{\partial q_i}, \quad i=1,\ldots,n \ ,$$

$$\frac{d^2q^{n+1}}{dt^2}=\frac{\partial V}{\partial q^i}\frac{dq^i}{dt}=-\frac{dL}{dt}, \quad \frac{d^2q^0}{dt^2}=0 \ . \tag{26}$$

The first n equations are obviously Newton's equations, the evolution equation for q^{n+1} is the differential version of Eq. (22), and the last equation states that $q^0=t$.

D. Geometrical description of average instability of the dynamics

Let us now briefly recall that, given an affine connection ∇ on a manifold M, the curvature tensor R of the connection is defined by [27]

$$R(X,Y)=\nabla_X\nabla_Y-\nabla_Y\nabla_X-\nabla_{[X,Y]} \tag{27}$$

with X, Y vector fields on M, and it measures the degree of noncommutativity of the covariant derivation ∇ on the manifold; if $R=0$ the manifold is flat and $[\nabla_X,\nabla_Y]=\nabla_{[X,Y]}$.

In local coordinates the components of the Riemann tensor are given by

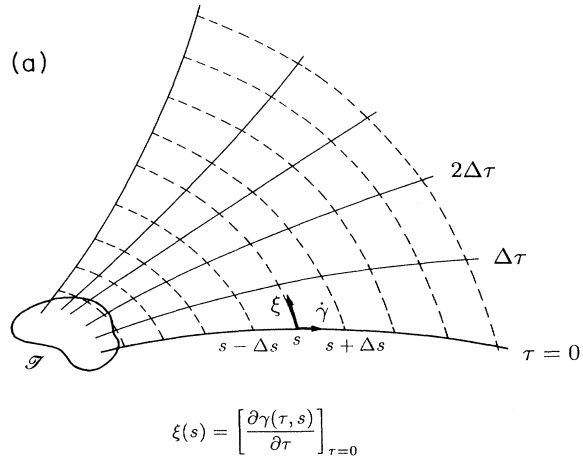
$$R^l_{kij}=\partial_i\Gamma^l_{kj}-\partial_j\Gamma^l_{ki}+\Gamma^m_{kj}\Gamma^l_{mi}-\Gamma^m_{ki}\Gamma^l_{mj} \ , \tag{28}$$

where $\partial_i=\partial/\partial q^i$.

The trace $R_{kl}=R^i_{kil}$ is the Ricci tensor and the scalar $R=g^{kl}R_{kl}$ is the scalar curvature of M.

It is important to remark that from the curvature properties of a Riemannian manifold some relevant consequences about the stability properties of its geodesics can be derived. This is the central point on which our attention is focused in the present work, so we comment on it in more detail.

Geodesics are obtained as extremals of a functional, but nothing more (i.e., if the extremals are maxima or minima) can be said unless second-order variations are considered. It is evident that these second-order variations give information about the surrounding landscape of a geodesic, thus of its stability with respect to small variations of the initial conditions. In order to measure this degree of sensitivity to initial conditions, we consider a congruence of geodesics $\gamma(\tau,s)$ and a separation vector ξ, which is Lie dragged by the congruence. Such a congruence, pictorially represented in Fig. 1(a), is the ensemble of geodesics issuing from a neighborhood $\mathcal{I}$ and locally filling the manifold without intersecting; moreover, it is assumed that this ensemble can be differentiably parametrized by a real number τ. Now choose any given geodesic $\gamma(\tau_0,s)$ as "reference geodesic," define the parametrization so that



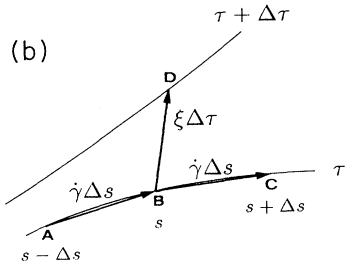


FIG. 1. Pictorial representation of a congruence of geodesics and of the geodesic separation vector ξ. (a) Bundle of geodesics, issuing from a neighborhood $\mathcal{I}$, parametrized by τ. (b) The separation vector can be used to measure the distance between nearby geodesics: $\xi\Delta\tau=\vec{BD}$ does from the point B of the reference geodesic $\gamma(s,\tau)$ to the point D of the test geodesic $\gamma(s,\tau+\Delta\tau)$. B and D correspond to the same value of the arc-length parameter s.

$\gamma(\tau_0,s)\equiv\gamma(\tau=0,s)$, then the separation vector $\xi(s)$ is defined by

$$\xi(s)=\left[\frac{\partial\gamma(\tau,s)}{\partial\tau}\right]_{\tau=0} \ .$$

Using finite differences in τ and s, Fig. 1(b) qualitatively shows that $\xi(s)$ can be used to approximate the distance between two nearby geodesics.

The Lie dragging by the flow of the separation vector is expressed by $\mathcal{L}_{\dot\gamma}\xi=0$ ($\mathcal{L}_{\dot\gamma}$ is the Lie derivative). Let us remember that $\nabla_{\dot\gamma}\dot\gamma(\tau,s)=0$ for any geodesic. Obviously the derivative $\nabla_{\dot\gamma}\xi$ measures the parallelism of nearby geodesics, while $\nabla_{\dot\gamma}\nabla_{\dot\gamma}\xi$ gives the acceleration of the separation ξ between initially close geodesics. The relationship between curvature and stability becomes evident by computing this acceleration. In fact, using $\mathcal{L}_XY=[X,Y]$ for the Lie derivative and the relation $\nabla_XY-\nabla_YX=[X,Y]$ for a connection with vanishing torsion, one has

$$\nabla_{\dot\gamma}\nabla_{\dot\gamma}\xi=\nabla_{\dot\gamma}(\mathcal{L}_{\dot\gamma}\xi+\nabla_{\xi}\dot\gamma)=\nabla_{\dot\gamma}\nabla_{\xi}\dot\gamma=[\nabla_{\dot\gamma},\nabla_{\xi}]\dot\gamma+\nabla_{\xi}\nabla_{\dot\gamma}\dot\gamma$$
$$=R(\dot\gamma,\xi)\dot\gamma \qquad (29)$$

and finally, using $R(X,Y)=-R(Y,X)$, one finds

$$\frac{\nabla}{ds}\frac{\nabla}{ds}\xi(s)+R_{\gamma(s)}(\xi(s),\dot\gamma(s))\dot\gamma(s)=0\ , \qquad (30)$$

which is the Jacobi–Levi-Civita equation [28] for a vector field of geodesic variations, where $(\nabla\xi/ds)=\nabla_{\dot\gamma}\xi$ is the covariant derivative along $\gamma(t)$; in local coordinates, it reads

$$\frac{\nabla\xi^i}{ds}=(\nabla_{\dot\gamma}\xi)^i=\frac{dx^j}{ds}\frac{\partial\xi^i}{\partial x^j}+\Gamma^i_{jk}\frac{dx^j}{ds}\xi^k=\frac{d\xi^i}{ds}+\Gamma^i_{ji}\frac{dx^j}{ds}\xi^k\ .$$

Equation (30) is our main tool for tackling Hamiltonian chaos from a geometrical point of view. It measures the sensitivity of the trajectories to initial conditions, and therefore their stability.

At this point one can wonder what is the difference with respect to the common definition of chaos, because here apparently the same definition is given. First of all, it is necessary to distinguish between the usual explanation of the origin of chaos and the operational method to detect it numerically (Lyapunov exponents). Dating back to Poincaré [2], Melnikov [3], and others, the origin of chaos is attributed to homoclinic intersections of perturbed separatrices near hyperbolic points. This is a perturbative picture. It requires the use of action-angle variables and applies to quasi-integrable Hamiltonians [Eq. (1)]. Moreover, for typical Hamiltonians of physical interest, the explicit change of coordinates to action-angle ones is, in general, a lengthy work that quickly becomes intractable with growing n. In practice, and particularly at high dimensionality, homoclinic intersections mainly give a *qualitative* picture of the source of chaos.

It has been already explained why we want to get rid of the limitations imposed by quasi-integrability. So, in the sake of a nontraditional explanation of the origin of Hamiltonian chaos and of a quantitative method to describe it, the central idea of this work is to exploit the tight relationship between *local instability* of the trajectories and *local* geometry (curvature) of the underlying manifold. The link between local instability and chaos is made by the compactedness of the manifolds where the trajectories live. The geometrical approach makes use of the natural coordinates (positions and velocities), applies at any energy and at any strength of the nonintegrable part of the Hamiltonian, and allows a unified treatment of both the explanation of the origin of chaos and the method to measure its intensity.

At variance, the standard operational way to detect chaos is through Lyapunov exponents: they are eigenvalues measuring the stability of nearby trajectories. Their mathematical definition is made possible by Oseledets's multiplicative theorem, which circumvents the problem of knowing the connection of the manifold where motion takes place.

Let us recall that for a generic flow Φ^t: $V\to V$ on a manifold V, given an invariant measure μ, and denoting by $d\Phi^t_x$: $T_xV\to T_{(\Phi^t x)}V$ its tangent dynamics, Oseledets's theorem ensures that $\forall x\in V_1$, $\forall e\in T_xV$ $(e\neq 0)$, with $V_1\subset V$ and such that $\mu(V_1)=1$, there exists and is finite the quantity [20]

$$\lambda(x,e)=\lim_{t\to\infty}\frac{1}{t}\ln\|d\Phi^t_x(e)\|\ , \qquad (31)$$

which is independent of the metric of V. The definition of the largest Lyapunov exponent relies upon this theorem. More explicitly, let $\underline{J}(x)$ be the Jacobian matrix of the Hamiltonian flow Φ^t at the point $x=(\mathbf{p},\mathbf{q})$. Denoting by ξ the variation vector, the tangent dynamics $d\Phi^t$ is described by

$$\frac{d\xi_i}{dt}=J_{ij}(x(t))\xi_k\ , \qquad (32)$$

thus $\underline{J}$: $T_xV\to T_{(\Phi^t x)}V$.

Take a vector $\xi_x\in T_xV$ and its transformed $\xi_{(\Phi^t x)}\in T_{(\Phi^t x)}V$, notice that

$$\langle\xi_{(\Phi^t x)},\underline{J}\xi_x\rangle=\langle\underline{J}^*\xi_{(\Phi^t x)},\xi_x\rangle\ , \qquad (33)$$

and that the product appearing at the left-hand side (lhs) belongs to $T_{(\Phi^t x)}V$, while the product of the right-hand side (rhs) necessarily belongs to T_xV, i.e., $\underline{J}^*\xi_{(\Phi^t x)}\in T_xV$, and in general $\|\xi_{(\Phi^t x)}\|\neq\|\xi_x\|$, whence $\underline{J}^*\neq\underline{J}^{-1}$; therefore $\underline{J}$ maps the tangent vectors forward in time while $\underline{J}^*$ maps these vectors backward in time but not retracing the forward evolution of $\xi(t)$. Having defined $\underline{G}=\prod_{i=1}^n\underline{J}_i(\Phi^{i-1}x)$, Oseledets's theorem states that the limiting matrix

$$\underline{\Lambda}_x=\lim_{n\to\infty}(\underline{G}^*\underline{G})^{1/2n} \qquad (34)$$

exists and is finite, thus for any n the product $(\prod_{i=1}^n\underline{J}_i^*)(\prod_{k=1}^n\underline{J}_k)$ maps an arbitrary vector $\xi^0_x\in T_xV$ into a vector $\xi^n_x\in T_xV$, i.e., into the *same* tangent space T_xV, hence an eigenvalue problem for $\underline{\Lambda}$ is well defined in the vector space T_xV.

Now it should be clear the standard Lyapunov exponents do not have any *local* meaning because they are defined as *asymptotic* quantities. In other words, it would be improper and misleading to think of Lyapunov exponents as averages of local divergence rates in phase space. An illuminating example of this fact is provided by a consequence of Fürstenberg's theorem on random matrices [29]. Consider an infinite product of matrices representing elliptic rotations with a random change of the rotation parameter (semiaxes of the ellipse), the limiting matrix has positive Lyapunov exponent, i.e., it represents a hyperbolic rotation. On the contrary, it is well evident that the average of the largest eigenvalue of the single matrices in the product (average of the local exponents) necessarily gives a vanishing Lyapunov exponent.

The above statements seem to be in contradiction with the standard numerical method used to estimate Lyapunov exponents. In fact, it is just an average of the local divergence rates of nearby trajectories that is numerically computed. But there is no mathematical proof—for generic flows—that true Lyapunov exponents

(those of Oseledets's theorem) are correctly approximated by this averaging procedure, whereas the standard numerical algorithm can be more naturally justified on different mathematical grounds: the Jacobi–Levi-Civita equation applied to the case of Eisenhart's metric (see below).

Equation (30) has an important property: it takes into account the connection on the manifold where natural motions take place, so it already contains information about parallel transport, i.e., how the reference frames are transported along a trajectory. At least for Hamiltonian flows, Eq. (30) makes possible a *local* description of the dynamics of a vector field of geodesic spread. Notice again that the ambient space of this intrinsic description is a manifold and not $\mathbb{R}^n$. Equation (30) allows a new definition of the dynamic instability exponents—analogous to Lyapunov exponents—without needing to go back to the initial tangent space in order to apply Oseledets's theorem.

Moreover, the instability exponents (i.e., chaoticity exponents for compact manifolds) that can be derived from Eq. (30) are directly related to the geometrical properties of the manifold underlying the motion, while this relationship remains quite implicit with Lyapunov exponents.

Finally, the use of *local* instead of *asymptotic* chaoticity indicators has another major consequence: their ensemble averages can be computed on constant-energy surfaces and compared with time averages computed along phase trajectories. On the contrary, this is not possible for Lyapunov exponents. They are sensitive to the history of a trajectory; only in case of strongly chaotic dynamics can some approximate computation be done [14], but this is limited to scaling behaviors.

Rewriting Eq. (30) in components, we get

$$\frac{\nabla}{ds}\frac{\nabla}{ds}\xi^i + R^i_{jlk}[\mathbf{q}(s)]\frac{dq^j}{dx}\xi^l\frac{dq^k}{ds}=0 \ . \tag{35}$$

Let us start with finding its explicit form in the case of Eisenhart's metric (23). Simple algebra shows that the only nonvanishing components of the Riemann tensor are given by ($a_{ij}=\delta_{ij}$)

$$R_{0i0j}=\frac{\partial^2 V}{\partial q^i \partial q^j} \ , \tag{36}$$

hence Eq. (35) becomes

$$\begin{aligned}
&\frac{\nabla}{ds}\frac{\nabla}{ds}\xi^i + R^i_{0j0}\left[\frac{dq^0}{ds}\right]^2\xi^j + R^i_{00j}\frac{dq^0}{ds}\xi^0\frac{dq^j}{ds} + R^i_{j00}\frac{dq^j}{ds}\xi^0\frac{dq^0}{ds}=0 \ , \\
&\frac{\nabla}{ds}\frac{\nabla}{ds}\xi^0 + R^0_{i0j}\frac{dq^i}{ds}\xi^0\frac{dq^j}{ds} + R^0_{0ij}\frac{dq^0}{ds}\xi^i\frac{dq^j}{ds}=0 \ , \\
&\frac{\nabla}{ds}\frac{\nabla}{ds}\xi^{n+1}=0 \ .
\end{aligned} \tag{37}$$

As $g^{00}=0$ [Eq. (28)], those components of the mixed Riemann tensor that have a contravariant index equal to 0 vanish, then since the parametrization of the arc length is affine we can replace s with t in ordinary and covariant derivatives. Finally, as ξ^0 does not accelerate, we choose separation vectors having the initial conditions $\xi^0(0)=\dot{\xi}^0(0)=0$, so that from Eq. (37) we get for the separation vector in configuration space

$$\frac{d^2\xi^i}{dt^2}+\left[\frac{\partial^2 V}{\partial q_i \partial q^j}\right]\xi^j=0, \quad i=1,\ldots,n \ , \tag{38}$$

which is exactly what is commonly used to compute numerical Lyapunov exponents. For Hamiltonian flows this equation is obviously equivalent to Eq. (32) since $\xi=(\xi_{\mathbf{q}},\xi_{\mathbf{p}})$ and $\xi_{\mathbf{p}}=d\xi_{\mathbf{q}}/dt$. If we denote by $((\mathbf{q}_\tau,\mathbf{p}_\tau),\ldots,(\mathbf{q}_{k\tau},\mathbf{p}_{k\tau}),\ldots)$ the numerical phase trajectory, then the variation vector $(\xi_{\mathbf{q}}(k\tau),\xi_{\mathbf{p}}(k\tau))$ is mapped to $(\xi_{\mathbf{q}}((k+1)\tau),\xi_{\mathbf{p}}((k+1)\tau))$ by the matrix

$$\underline{J}_k=\begin{bmatrix}\underline{1} & \tau\underline{1}\\ \tau\underline{\Omega} & \underline{1}+\tau^2\underline{\Omega}\end{bmatrix}_{\mathbf{q}(k\tau)} \ , \tag{39}$$

where $\underline{\Omega}=-\mathcal{H}(V)$ is the Hessian of V, so that a numerical estimate of true Lyapunov exponents would first require the computation of the product matrix $(\prod_{i=1}^n \underline{J}_i^*)(\prod_{k=1}^n \underline{J}_k)$ with a large n [Eq. (34)] and then its diagonalization would give Lyapunov characteristic exponents. On the contrary, the universally adopted [30] numerical algorithm prescribes the averaging of *local* divergence rates of nearby trajectories: this procedure is meaningful and well defined in the light of the geometrical background of Eq. (38) and does not require Oseledets's theorem at all. Moreover, the average of local divergence rates is insensitive to the order of the addenda, while the product of the matrices $\underline{J}_k$ is very sensitive to their order because they do not commute (Fürstenberg's theorem is a consequence of noncommutativity).

Let us now consider the case of a kinetic-energy metric on M. Having set $g_{ij}=W\delta_{ij}$ with δ_{ij} the Kronecker δ function and $W=E-V(q^a)$, using Eq. (7) we find

$$\Gamma^h_{ij}=\frac{1}{E-V}\delta^{hm}(\partial_i W\delta_{jm}+\partial_j W\delta_{mi}-\partial_m W\delta_{ij}) \tag{40}$$

for the connection coefficients, and according to Eq. (28) the components of the Riemann tensor are

$$R_{ijkl}=\frac{1}{2}\left[-\frac{\partial^2 V}{\partial q^j \partial q^k}\delta_{il}+\frac{\partial^2 V}{\partial q^i \partial q^k}\delta_{jl}+\frac{\partial^2 V}{\partial q^j \partial q^l}\delta_{ik}-\frac{\partial^2 V}{\partial q^i \partial q^l}\delta_{jk}\right]$$
$$-\frac{3}{4(E-V)}\left[\frac{\partial V}{\partial q^j}\frac{\partial V}{\partial q^k}\delta_{il}-\frac{\partial V}{\partial q^i}\frac{\partial V}{\partial q^k}\delta_{jl}-\frac{\partial V}{\partial q^j}\frac{\partial V}{\partial q^l}\delta_{ik}+\frac{\partial V}{\partial q^i}\frac{\partial V}{\partial q^l}\delta_{jk}\right]+\frac{1}{4(E-V)}(\delta_{jk}\delta_{il}-\delta_{ik}\delta_{jl})\frac{\partial V}{\partial q^h}\frac{\partial V}{\partial q^h}\,. \tag{41}$$

Hence, if we define an operator $\underline{Q}$ whose matrix elements are

$$Q^i_l(\mathbf{q}(s),\dot{\mathbf{q}}(s))=\frac{1}{2W^2}g^{im}R_{mjlk}\left[\frac{dq^j}{dt}\right]\left[\frac{dq^k}{dt}\right], \tag{42}$$

then Eq. (35) becomes

$$\frac{\nabla}{ds}\frac{\nabla}{ds}\xi^i+Q^i_l(\mathbf{q}(s),\dot{\mathbf{q}}(s))\xi^l=0\,, \tag{43}$$

so that the n eigenvalues of the matrix $\underline{Q}$ can be used to define n independent instability exponents. Notice that, in case of practical evaluation of such exponents, the second-order covariant derivative—written in coordinates—yields additional terms modifying the matrix $\underline{Q}$.

A negative eigenvalue indicates the runaway acceleration of the corresponding component of the separation vector. The eigenvalues can be averaged along a trajectory to give, for compact manifolds, n geometric indicators of chaoticity.

It is well evident that this is a hard numerical work because there are $\sim n^4$ independent components of the Riemann tensor. Therefore for the sake of the simplification of the problem, let us begin by studying the time evolution of the norm $\|\xi\|$ of the separation vector looking for some synthetic indicator of chaos similar to the largest Lyapunov exponent (numerical).

To this purpose let us multiply Eq. (30) by ξ,

$$\langle\nabla^2_{\dot\gamma}\xi,\xi\rangle=-\langle R(\xi,\dot\gamma)\dot\gamma,\xi\rangle\,, \tag{44}$$

where $\langle\ ,\ \rangle$ stands for the scalar product on the manifold (in local coordinates $\langle X,Y\rangle=g_{ij}X^iY^j$).

The lhs of Eq. (44) can be simplified by considering the identity

$$\langle\nabla^2_{\dot\gamma}\xi,\xi\rangle=\nabla_{\dot\gamma}\langle\nabla_{\dot\gamma}\xi,\xi\rangle-\|\nabla_{\dot\gamma}\xi\|^2 \tag{45}$$

and the inequality

$$\|\nabla_{\dot\gamma}\xi\|^2\geq\left[\frac{d}{ds}\|\xi\|\right]^2,$$

so that from

$$\nabla_{\dot\gamma}\langle\nabla_{\dot\gamma}\xi,\xi\rangle=\frac{1}{2}\frac{d^2}{ds^2}\|\xi\|^2$$

we get

$$\langle\nabla^2_{\dot\gamma}\xi,\xi\rangle\leq\frac{1}{2}\frac{d^2}{ds^2}\|\xi\|^2-\left[\frac{d}{ds}\|\xi\|\right]^2. \tag{46}$$

For what concerns the rhs of Eq. (44), remember that for any given couple of vectors $X,Y\in T_pM$ (where $p\in M$) and nondegenerate plane $\pi(p)\in T_pM$ spanned by them, the function on T_pM defined by

$$K(X,Y)=\frac{\langle R(X,Y)X,Y\rangle}{\langle X,X\rangle\langle Y,Y\rangle-\langle X,Y\rangle^2} \tag{47}$$

is the so-called sectional curvature at p. Therefore the rhs of Eq. (44) can be rewritten as

$$\langle R(\xi,\dot\gamma)\dot\gamma,\xi\rangle=K(\dot\gamma,\xi)(\langle\dot\gamma,\dot\gamma\rangle\langle\xi,\xi\rangle-\langle\dot\gamma,\xi\rangle^2)\,, \tag{48}$$

and supposing that ξ denotes only the normal components to $\dot\gamma$, Eq. (48) becomes

$$\langle R(\xi,\dot\gamma)\dot\gamma,\xi\rangle=K(\dot\gamma,\xi)\langle\dot\gamma,\dot\gamma\rangle\langle\xi,\xi\rangle\,. \tag{49}$$

At this point some approximation must be introduced to simplify the original problem, because the sectional curvature at $p\in M$ still depends on the full Riemann tensor and on the vectors $\dot\gamma,\xi\in T_pM$; thus we try to replace $K(\dot\gamma,\xi)$ by some average quantity at $p\in M$. We have at least two natural possibilities at our disposal. In fact, if $X_{(r)}$ and $X_{(s)}$ are mutually orthogonal unit vectors of T_pM, all the possible sectional curvatures at $p\in M$ are given by

$$K^{(2)}_{rs}=R_{ijlk}X^i_{(r)}X^j_{(s)}X^l_{(r)}X^k_{(s)},\quad r,s=1,\ldots,n\,, \tag{50}$$

where n is the dimension of the manifold, then it is well known [31] that (i) $\sum_{r,s=1}^n K^{(2)}_{rs}=R$, with R the scalar curvature of M at p and (ii) $\sum_{s=1}^n K^{(2)}_{rs}=\Re(X_{(r)})=R_{ik}X^i_{(r)}X^k_{(r)}$, where $\Re(X_{(r)})$ is the Ricci curvature in the direction of the unit tangent vector $X_{(r)}$.

Take $X_{(r)}$ and $X_{(s)}$ codirectional with $\dot\gamma$ and ξ, respectively. There exist n independent directions for $\dot\gamma$ at p and $n-1$ for ξ such that $\xi\perp\dot\gamma$. Let us average over all the combinations of $\dot\gamma_p$ and ξ_p, thus over all the geodesics issuing from $p\in M$ and having $\dot\gamma_p$ as an initial condition, and over all the different choices for the geodesic deviation at p (take, for instance, a uniform distribution of directions for ξ keeping its norm fixed). Now $K(\dot\gamma,\xi)$ in Eq. (49) is replaced by the average sectional curvature $K_{\rm av}=R/n(n-1)$, so that the rhs of Eq. (44) becomes

$$K_{\rm av}\left\|\frac{d\gamma}{ds}\right\|^2\|\xi\|^2\equiv\frac{R}{n(n-1)}\|\xi\|^2\,, \tag{51}$$

since $\|d\gamma/ds\|^2=1$. A less drastic approximation consists in choosing a given geodesic that originates at p, i.e., holding $\dot\gamma_p$ fixed, and averaging only over ξ_p. In this case, after the above item (ii), we replace $K(\dot\gamma,\xi)$ in Eq. (49) by $\Re(d\gamma/ds)$, so that now the rhs of Eq. (44) simplifies to

$$\frac{1}{n}\Re\left[\frac{d\gamma}{ds}\right]\|\xi\|^2=\frac{1}{n}R_{ik}\frac{d\gamma^i}{ds}\frac{d\gamma^k}{ds}\|\xi\|^2$$
$$\equiv\frac{1}{n}R_{ik}\frac{dq^i}{ds}\frac{dq^k}{ds}\|\xi\|^2 . \quad (52)$$

Let us now give explicit expressions for the Ricci tensor and for the scalar curvature. By contracting the Riemann tensor we obtain

$$R_{ik}=\left[\frac{\Delta V}{2(E-V)}+\frac{4-n}{4(E-V)^2}(\nabla V)^2\right]\delta_{ik}$$
$$+\frac{n-2}{2(E-V)}\frac{\partial^2 V}{\partial q^i\partial q^k}+\frac{3(n-2)}{4(E-V)^2}\frac{\partial V}{\partial q^i}\frac{\partial V}{\partial q^k} \quad (53)$$

and then the trace R_i^i gives the scalar curvature

$$R=n(n-1)\left[\frac{\Delta V}{n(E-V)^2}-\left[\frac{1}{4}-\frac{3}{2n}\right]\frac{(\nabla V)^2}{(E-V)^3}\right], \quad (54)$$

where Δ and ∇ stand for the Euclidean Laplacian and gradient, respectively.

Finally, denote by χ either the quantity in Eq. (51) or that in Eq. (52), consider $\|\xi\|^2$ as a corresponding averaged norm, then our approximate version of Eq. (44) reads as

$$\frac{d^2\|\xi\|^2}{ds^2}+2\hat{\chi}\|\xi\|^2-2\left[\frac{d\|\xi\|}{ds}\right]^2\geq 0 , \quad (55)$$

$$\hat{\chi}=R/n(n-1) , \quad (55a)$$

$$\hat{\chi}=\frac{1}{n}R_{ik}\frac{dq^i}{ds}\frac{dq^k}{ds} , \quad (55b)$$

which is the main computational tool of the present paper. The quantities appearing in items (55a) and (55b) are, respectively, given by Eqs. (54) and (53), according to the corresponding approximations, and are computed along a trajectory.

In what follows we denote by ζ the average norm $\|\xi\|^2$ of the separation vector, therefore Eq. (55) is rewritten as

$$\frac{d^2\zeta}{ds^2}+2\hat{\chi}\zeta-\frac{1}{2\zeta}\left[\frac{d\zeta}{ds}\right]^2\geq 0 . \quad (56)$$

In order to describe the time evolution of ζ, i.e., passing from geometry to dynamics, we make use of the arc-length parametrization $ds^2=2W^2dt^2$ to find

$$\frac{d^2\zeta}{dt^2}-\frac{1}{W}\frac{dW}{dt}\frac{d\zeta}{dt}+2\chi\zeta-\frac{1}{2\zeta}\left[\frac{d\zeta}{dt}\right]^2\geq 0 . \quad (57)$$

Now $\chi=2RW^2/n(n-1)$ or, alternatively, $\chi=R_{ik}\dot{q}^i\dot{q}^k/n$. The associated equation (i.e., with the lhs strictly equal to zero) is homogeneous of degree 1 in $\ddot{\zeta},\dot{\zeta},\zeta$, so that—in principle—it can be solved by substituting

$$\zeta(t)=\zeta_0\exp\int f(t)dt \quad (58)$$

into Eq. (57) and then integrating the resulting equation for $f(t)$

$$\frac{df}{dt}+\frac{1}{2}f^2-\left[\frac{1}{W}\frac{dW}{dt}\right]f+2\chi=0 . \quad (59)$$

Equation (59) is a Riccati equation. The functions $-\dot{W}(t)/W(t)$ and $\chi(t)$ are rapidly oscillating functions with complicated time behaviors. By the standard transformation

$$f=\frac{2}{x}\frac{dx}{dt} , \quad (60)$$

Eq. (59) becomes a linear second-order equation with nonconstant coefficients

$$\frac{d^2x}{dt^2}-\frac{\dot{W}(t)}{W(t)}\frac{dx}{dt}+\chi(t)x=0 . \quad (61)$$

Finally, by introducing the transformation

$$y=x\exp\left[-\tfrac{1}{2}\int dt\dot{W}(t)/W(t)\right] , \quad (62)$$

we rewrite Eq. (61) as

$$\frac{d^2y}{dt^2}+Q(t)y=0 , \quad (63)$$

where

$$Q(t)=\chi(t)-\frac{1}{4}\left[\frac{\dot{W}}{W}\right]^2+\frac{1}{2}\frac{d}{dt}\left[\frac{\dot{W}}{W}\right] . \quad (64)$$

Equation (63) is called Hill's equation.

As we shall see in the next section, the numerical computation at any energy of $\dot{W}/W$ shows that this is a rapidly oscillating function of zero mean value and very small norm, whereas $\chi(t)=\bar{\chi}+\eta(t)$ has a nonzero mean $\bar{\chi}$ and a fluctuating part $\eta(t)$. As a consequence the exponential in Eq. (62) makes small oscillations around 1, and $y(t)$ is close to $x(t)$.

It is well known that parametric resonance can make the solutions of Hill's equation unstable. This is a subtle mechanism to make chaos [32], in fact, even if $\chi(t)$ is always positive, provided that the conditions for parametric instability are satisfied, $y(t)$ can grow exponentially in time and thus, going through Eqs. (60) and (58), the norm ζ of the geodesic variation vector also grows exponentially in time.

In the next section it is shown that also this mechanism of parametric instability is actually at work to make chaos. However a systematic and quantitative investigation of this aspect is left for a subsequent work. Let us keep in mind that Eq. (63) has nontrivial instability properties and deserves a more complete study; however, we shall only proceed with a first rough stability analysis.

$Q(t)$ is rapidly oscillating with respect to the instability time scales (measured, for instance, by the numerical Lyapunov exponents), thus we can perform a simple averaging of Eq. (63). The averaged *reference* system becomes

$$\frac{d^2\bar{y}^{(\pm)}}{dt^2}+\bar{Q}\bar{y}^{(\pm)}=0,\quad \bar{Q}=\frac{1}{T}\int_0^T dtQ(t)\simeq\bar{\chi},\quad T \text{ large} . \quad (65)$$

The suffix $(\pm)$ refers to the solutions $\bar{y}^{(\pm)}(t)$ corresponding to positive or negative $\bar{Q}$. Neither Eq. (65) nor Eq. (63) has singularities at finite times, and Eq. (65) is a uniform mean of Eq. (63), hence the conditions required by a theorem due to Bogoliubov [33], are fulfilled, and so the smallness of the norm $\|y^{(\pm)}(t)-\bar{y}^{(\pm)}(t)\|$ is ensured. In other words, during the time intervals where $\chi(t)<0$, the solution $y^{(-)}(t)$ of Eq. (63) is close to the averaged solution $\bar{y}^{(-)}(t)$ of Eq. (65), where $\bar{\chi}$ equates the local negative average of $\chi(t)$. Obviously it is $\bar{y}^{(-)}(t)=A\cosh(\sqrt{-\bar{\chi}}t)+B\sinh(\sqrt{-\bar{\chi}}t)$.

Conversely, when $\chi(t)>0$, the solution $y^{(+)}(t)$ of Eq. (63) is close to the averaged solution of Eq. (65): $\bar{y}^{(+)}(t)=A\cos(\sqrt{\bar{\chi}}t)+B\sin(\sqrt{\bar{\chi}}t)$. By substituting $x^{(\pm)}(t)\simeq y^{(\pm)}(t)$ into Eq. (60) we find $f(t)$, and through Eq. (58) we finally find an approximate solution of Eq. (57).

In the case $\chi(t)>0$, the solution of Eq. (57) is approximated by

$$\zeta(t)\approx\zeta_0\exp\int dt\left[2\frac{-\sqrt{\bar{\chi}}\sin(\sqrt{\bar{\chi}}t)+B\sqrt{\bar{\chi}}\cos(\sqrt{\bar{\chi}}t)}{\cos(\sqrt{\bar{\chi}}t)+B\sin(\sqrt{\bar{\chi}}t)}\right]$$
$$=\zeta_0[\cos(\sqrt{\bar{\chi}}t)+B\sin(\sqrt{\bar{\chi}}t)]^2\,, \tag{66}$$

where $A=1$ has been chosen to make $\zeta(0)=\zeta_0$, and B is an arbitrary parameter. It is well evident that $\chi(t)>0$ implies that the norm of the separation vector can remain bounded, so that initially close trajectories remain close to one another. Notice that this is a *necessary* condition for the occurrence of regular motion but it is not sufficient because of the above-mentioned possibility of parametric instability.

At variance, in the case $\chi(t)<0$ it is found

$$\zeta(t)\approx\zeta_0\exp\int dt\left[2\frac{\sqrt{-\bar{\chi}}\sinh(\sqrt{-\bar{\chi}}t)+B\sqrt{-\bar{\chi}}\cosh(\sqrt{-\bar{\chi}}t)}{\cosh(\sqrt{-\bar{\chi}}t)+B\sinh(\sqrt{-\bar{\chi}}t)}\right]$$
$$=\zeta_0[\cosh(\sqrt{-\bar{\chi}}t)+B\sinh(\sqrt{-\bar{\chi}}t)]^2\sim\tfrac{1}{4}\zeta_0(1+B^2)\exp(2\sqrt{-\bar{\chi}}t)\,, \tag{67}$$

again $A=1$ is chosen to make $\zeta(0)=\zeta_0$ and B is arbitrary. Therefore $\chi(t)<0$ is a *sufficient* condition for local exponential instability of nearby trajectories or, in other words, for chaos (on compact manifolds). This also means that the natural motion makes stationary *but not minimum* the action—or arc-length—functional and this fact is responsible for the local defocalization of the geodesics, thus for their extreme sensitivity to initial conditions.

It is natural to take $\Lambda=\mathrm{Re}\sqrt{-\chi}$ as a *geometric-chaoticity indicator* (GCI). In order to measure a *global* degree of chaoticity of the trajectories, we can use time averages of Λ; these are computed along any natural motion of the system under investigation as

$$\bar{\Lambda}^T=\frac{1}{T}\int_0^T dt\,\mathrm{Re}[-\chi(q^a(t),\dot{q}^a(t))]^{1/2}\,, \tag{68}$$

and can be obtained only by numerical integration of the equations of motion.

We can also define a *static* average $\langle\Lambda\rangle$ by

$$\langle\Lambda\rangle=\Omega^{-1}\int_{\Sigma_E}d\sigma_E\,\mathrm{Re}[-\chi(q^a,\dot{q}^a)]^{1/2}$$
$$=\Omega^{-1}\int\prod_{i=1}^{n}dq_i d\dot{q}_i\,\delta(H'(q^a,\dot{q}^a)-E)$$
$$\times\mathrm{Re}[-\chi(q^a,\dot{q}^a)]^{1/2}\,, \tag{69}$$

where Σ_E is the constant-energy surface of phase space defined by $H'(q^a,\dot{q}^a)=E$, and $H'(q^a,\dot{q}^a)$ is the Hamiltonian of the system, with p^a replaced by $\dot{q}^a$ (for hyper-regular Lagrangians TM and T^*M are fully interchangeable). In this way we compute an average quantity that retains some geometrical property of the manifold (M,g_J) where the dynamics takes place. Using an observable of geometric type, one can wonder whether the transition between weak and strong chaoticity can be recovered by following a corresponding change in the geometric properties of (M,g_J). This is the main concern of the following section.

III. NUMERICAL COMPUTATIONS

In what follows we report the results of numerical computations of dynamic and static averages of Λ or other geometrical quantities for three model Hamiltonians. These are

$$H(\mathbf{p},\mathbf{q})=\sum_{i=1}^{n}[\tfrac{1}{2}p_i^2+\tfrac{1}{2}(q_{i+1}-q_i)^2+\tfrac{1}{4}\mu(q_{i+1}-q_i)^4] \tag{70}$$

for the FPU β model;

$$H(\mathbf{p},\mathbf{q})=\sum_{i=1}^{n}[\tfrac{1}{2}p_i^2+\tfrac{1}{2}(q_{i+1}-q_i)^2+\tfrac{1}{2}m^2q_i^2+\tfrac{1}{4}\mu q_i^4] \tag{71}$$

for the so-called lattice φ^4 model, and

$$H(\mathbf{p},\mathbf{q})=\sum_{i=1}^{n}\left[\tfrac{1}{2}p_i^2+\frac{1}{b}(e^{-b(q_{i+1}-q_i)}-1)\right] \tag{72}$$

for the Toda lattice. We remind the reader that the Toda lattice is an integrable model.

The equations of motion derived from (70), (71), and (72) have the form

$$\frac{d^2q_i}{dt^2}=(q_{i+1}+q_{i-1}-2q_i)+F_i(\mathbf{q}(t))\equiv G_i(\mathbf{q}(t))\,, \tag{73}$$

where

$$F_i(\mathbf{q}(t))=\begin{cases}-m^2q_i-\mu q_i^3, & \varphi^4\text{ model}\\ \mu[(q_{i+1}-q_i)^3-(q_i-q_{i-1})^3], & \beta\text{-FPU model}\,,\end{cases}$$
$$G_i(\mathbf{q}(t))=-ae^{-b(q_{i+1}-q_i)}+ae^{-b(q_i-q_{i-1})},\ \text{Toda lattice}\,. \tag{74}$$

Boundary conditions have been chosen as periodic, that is, $q_i = q_{n+i}$.

Equations obtained from (73) and (74) have been numerically integrated by means of a leap-frog algorithm, which is an explicit scheme given by

$$q_i(t+\Delta t) = 2q_i(t) - q_i(t-\Delta t) + (\Delta t)^2 G_i(\mathbf{q}(t)) \ , \tag{75}$$

where the truncation error is $O[(\Delta t)^4]$. Besides its simplicity, this algorithm has the great advantage of being symplectic, which means that its effect upon the equations of motion is equivalent to a canonical transformation of variables. This ensures a faithful numerical representation of a Hamiltonian flow.

All the models have been integrated at $n = 128$ in order to compare the results of the present work with those reported in Ref. [14]. For the same reason, the coupling constant μ in Eq. (70) is equal to 0.1, and the constants m^2 and μ in Eq. (71) are, respectively, equal to 0.01 and 0.1. The choice of the constants in Eq. (72) is $a = 10$ and $b = 0.1$; with these values the second-order expansion of the exponential coincides with the harmonic-oscillator part of the other models.

Numerical simulations have been performed on a Cray Y-MP computer working, in single precision, with words of 64 bits. The dynamics is numerically computed using an integration time step $\Delta t = 0.01$; this value is reduced at large energies in order to keep the relative energy fluctuations $\Delta E/E$ in the interval 10^{-6}–10^{-5}. We recall that the leap-frog algorithm gives no energy drift in time, i.e., the energy fluctuations have zero mean.

Initial conditions are chosen at random and at equipartition among the harmonic modes—as in Ref. [14]—for the three models, that is $q_i(0) = 0$ and $p_i(0) = \dot{q}_i(0) = p_i^{(0)}$, with

$$p_i^{(0)} = \left[\frac{2}{n}\right]^{1/2} \sum_{j=k_1,\ldots,k_{n/2}} \left[\dot{A}_j(0) \cos\left[\frac{2\pi}{n} ij\right] + \dot{B}_j(0) \sin\left[\frac{2\pi}{n} ij\right] \right] , \tag{76}$$

where $\dot{A}_i(0)$ and $\dot{B}_i(0)$ are randomly selected and constrained by

$$\begin{aligned} &\dot{A}_{k_j}^2(0) + \dot{B}_{k_j}^2(0) = \frac{4E_0}{n}, \quad k_j = 1, \ldots, \frac{n}{2} - 1 \\ &\dot{A}_{k_j}^2(0) = \frac{2E_0}{n}, \quad k_j = 0, \frac{n}{2} \end{aligned} \tag{77}$$

in order to fix at E_0 the energy of the initial condition. The integration algorithm (75) is initialized by $q_i(0) = 0$ and $q_i(-\Delta t) = -p_i^{(0)} \Delta t$.

For both the FPU model and the Toda lattice, particular care has been paid to prevent any possible drift of the center of mass due to numerical errors. It has been also checked that all the q_i remain limited during their time evolution. A slightly modified version of the FPU model has been also considered with a small "mass term," i.e., $\frac{1}{2} m^2 q_i^2$ with $m^2 = 10^{-6}$ in the potential, in order to verify that no qualitative change was produced on the results reported in the following. All these checks aimed at getting rid of the possible doubts about compactedness of the configuration-space manifold. Moreover, we recall that the q_i in the models under investigation represent physical displacements of atoms from their respective equilibrium positions in a chain, therefore they cannot indefinitely grow.

At each integration step the scalar curvature R of M is computed using Eq. (54). For the FPU model, Fig. 2(a) shows the number of times $N^{(-)}$ that R is found negative

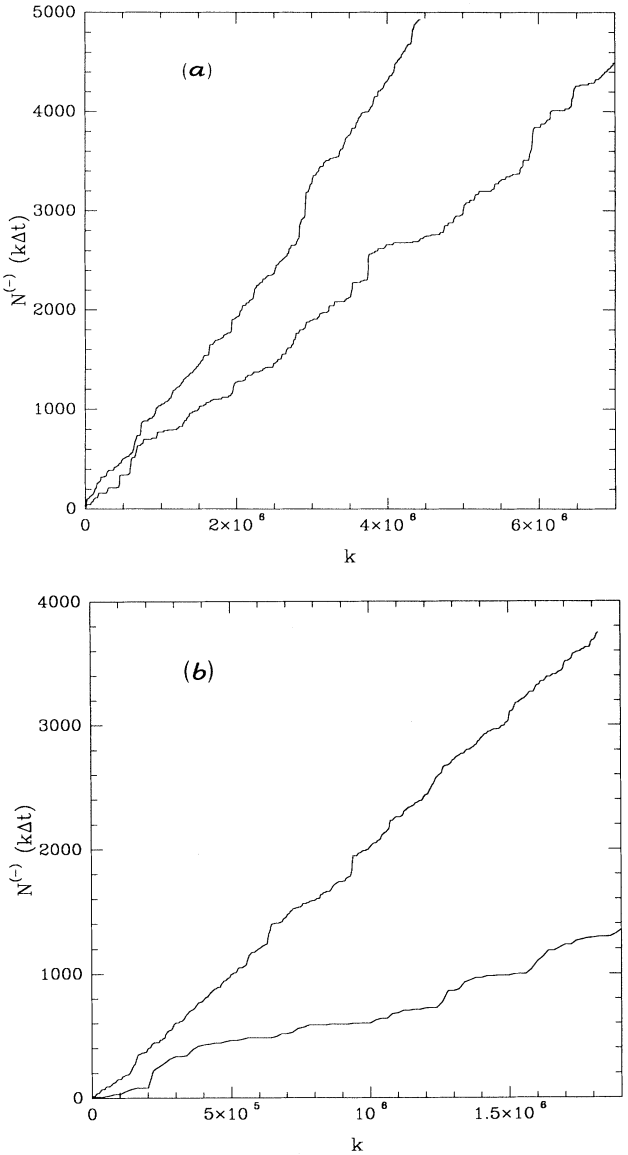


FIG. 2. FPU model, $n = 128$. (a) Number of times that the scalar curvature is found negative vs the number of integration steps k. Upper curve at $\ln\varepsilon = 4$; lower curve at $\ln\varepsilon = 2$. (b) Number of times that the Ricci curvature if found negative vs the number of integration steps k. Upper curve at $\ln\varepsilon = 4$; lower curve at $\ln\varepsilon = 1.25$. All the results are scaled to account for the different integration time steps.

as a function of the number of integration steps k. It is well evident that $N^{(-)}(k)$ looks like a staircase function linearly growing in time, in the average.

A definitely similar result is found in the case of the φ^4 model. On the contrary, in the case of the Toda lattice and in the harmonic limits ($\mu=0$) of Eqs. (70) and (71), R is *always* found *positive.*

Qualitatively analogous results are obtained for $\Re(\dot{\mathbf{q}})=R_{ik}(q^a)\dot{q}^i\dot{q}^k<0$. In Fig. 2(b) the number of times that $\Re(\dot{\mathbf{q}})<0$ is shown for different values of the energy density ε. As in the previous case, a staircase function shows up that, in the average, is a linearly growing function of time with a slope that changes with ε. The results concern the FPU model and, again, things are quite similar for the φ^4 model.

In the harmonic limits of these models, as well as in the case of the Toda lattice, $\Re(\dot{\mathbf{q}})>0$ is *always* found. In the case of the Toda lattice, if b is not much smaller than 1, some additional caution is necessary in the numerical integration at large ε. Double precision (128 bits per word) and very small integration time steps are necessary to avoid numerical artifacts that break the integrability of the model.

These results are the necessary prerequisites in view of a consistent differential—geometrical description of chaos. In fact, in the integrable cases, χ is found *always positive,* and no parametric instability is detected (see below), thus the geodesics of (M,g_J) are stable with respect to small variations of the initial conditions and chaos is absent. At variance, for the FPU and φ^4 models that are not integrable, $R<0$ and $\Re(\dot{\mathbf{q}})<0$ are found in subsets of a nonvanishing measure of M; according to Eqs. (57) and (67) this is sufficient to make chaos.

It is worth noticing that, in order to have a nonvanishing probability of numerically picking a truly regular trajectory, a positive measure of such trajectories must exist in phase space, which is possible only below the KAM threshold. Assuming the n dependence of this threshold as given in the Introduction, and disregarding finer details just to get a rough estimate of μ_c, at $n=128$ we find $\mu_c\sim 10^{-270}$. This is an exceedingly small perturbation strength, so it is hopeless to distinguish numerically between *truly* regular and stochastic orbits in the same model. At variance, regular trajectories can be studied in integrable models.

Both Ricci curvature $\Re(\dot{\mathbf{q}}(t))$ and scalar curvature $R(\dot{q}(t))$ are strongly fluctuating functions along any trajectory in FPU and φ^4 models. In Fig. 3 a typical example is given of the scalar curvature for the FPU model. The Ricci curvature has a very similar time behavior. This fact suggests that (M,g_J) has a high degree of "bumpiness."

The quantity $w(t)=(1/W)(dW/dt)$ entering Eqs. (61) and (64) rapidly oscillates with zero mean, and with a root-mean-square value that is weakly dependent on ε. For instance, at $E=4239$ ($\ln\varepsilon=3.5$) the average value $\overline{w}^2$ is 0.014 and at $E=47$ ($\ln\varepsilon=-1$) it is 0.007. Correspondingly, when χ stands for the Ricci curvature, at $E=4239$ it is $\overline{w}^2\simeq 0.012\chi$ and at $E=47$ it is $\overline{w}^2\simeq 0.04\chi$.

As already discussed in the previous section, an oscillatory behavior of $\chi(t)$ could make chaos through parametric resonance, and this may happen also if $\chi(t)$ is always positive. This possibility has been checked by numerically integrating Hill's equation (63) together with the equations of motion of the FPU model. The results show that this mechanism is actually at work.

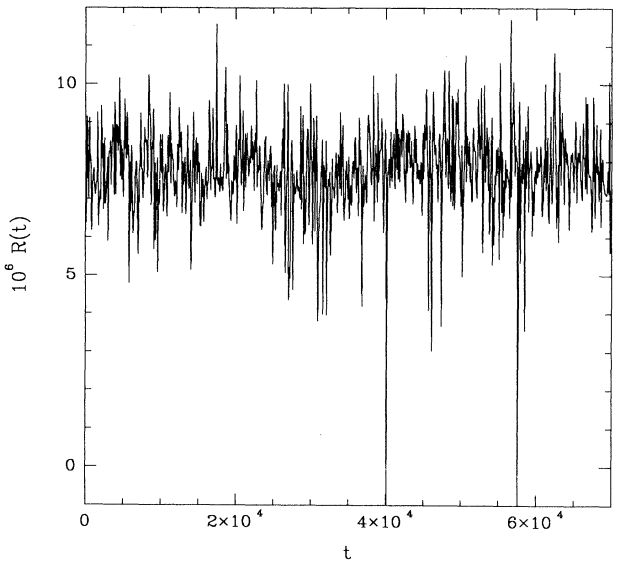


FIG. 3. FPU model. Scalar curvature vs time at $\ln\varepsilon=2$.

An example of this phenomenon at $E=2570$ ($\ln\varepsilon=3$) is shown in Fig. 4(a). The numerical solution $y(t)$ of Eq. (63) has been obtained using $\chi(t)=(1/n)R_{ik}(t)\dot{q}^i(t)\dot{q}^k(t)$ and a leap-frog algorithm with $\Delta t=0.001$. The solution $y(t)$ oscillates with an increasing amplitude. This amplitude growth can be due to negative values of $\chi(t)$ or to parametric resonance. In order to separate the effect due to parametric instability, all the negative values of $\chi(t)$ have been replaced by zero during the integration of Eq. (63). A test—performed by keeping also the negative values of $\chi(t)$—showed that the solution $y(t)$ is not significantly altered. It is very interesting to compare the result of Fig. 4(a) with the solution $y(t)$ obtained in the harmonic limit ($\mu=0$) at the same energy $E=2570$, plotted in Fig. 4(b). In fact, also at $\mu=0$ $\chi(t)$ appears strongly fluctuating, however in this case $y(t)$ makes only bounded oscillations. In Fig. 4(c) a synopsis of both results is given in log-lin scale. In particular, the average exponential growth of the envelope of $y(t)$ is well evident. A systematic investigation about the generation of chaos by parametric instability will be the subject of a future work.

Let us now comment about the ε dependence of GCI. In Fig. 5 we report some examples of the time behavior of

$$\overline{\Lambda}(t)=\frac{1}{t}\int_0^t dt'\,\mathrm{Re}\left[-\frac{1}{n}R_{ik}(q^a(t'))\dot{q}^i(t')\dot{q}^k(t')\right]^{1/2} \tag{78}$$

computed for the FPU model and using Eq. (53) for the Ricci tensor.

The convergence rate is in general rather good; howev-

er, the numerical computations are very time consuming. In particular, at small values of ε, the computer time needed gets rather long: the point at $\ln\varepsilon = -2$ in Fig. 6(a), for example, needed 6 h of CPU time on a Cray Y-MP computer.

Figure 6(a) shows the result obtained for $\Lambda(\varepsilon)$. We denote by Λ the stabilized ("asymptotic") value of $\bar{\Lambda}(t)$. A synopsis is also given of the ε dependence of the largest numerical Lyapunov exponent λ_1, already reported in Ref. [14], through which the SST was defined. The very interesting and nontrivial result is the close analogy between $\lambda_1(\varepsilon)$ and $\Lambda(\varepsilon)$, and the coincidence—within the numerical precision—of the critical values of ε at which their crossovers occur.

At small ε the computation of Λ is not extended as far as λ_1, the reason being merely practical. The lowering of ε below $\ln\varepsilon = -2$ yields a too steep increase of the computational time.

The average (78) has been also computed by replacing $\chi = 2RW^2/n(n-1)$, but in this case $\bar{\Lambda}(t)$ has stronger fluctuations and a definitely worse convergence in time. However, the final result, say $\Lambda_S(\varepsilon)$, roughly displays a transitional behavior around the same critical energy density. $\Lambda(\varepsilon)$ and $\Lambda_S(\varepsilon)$ are compared in Fig. 6(b). It is

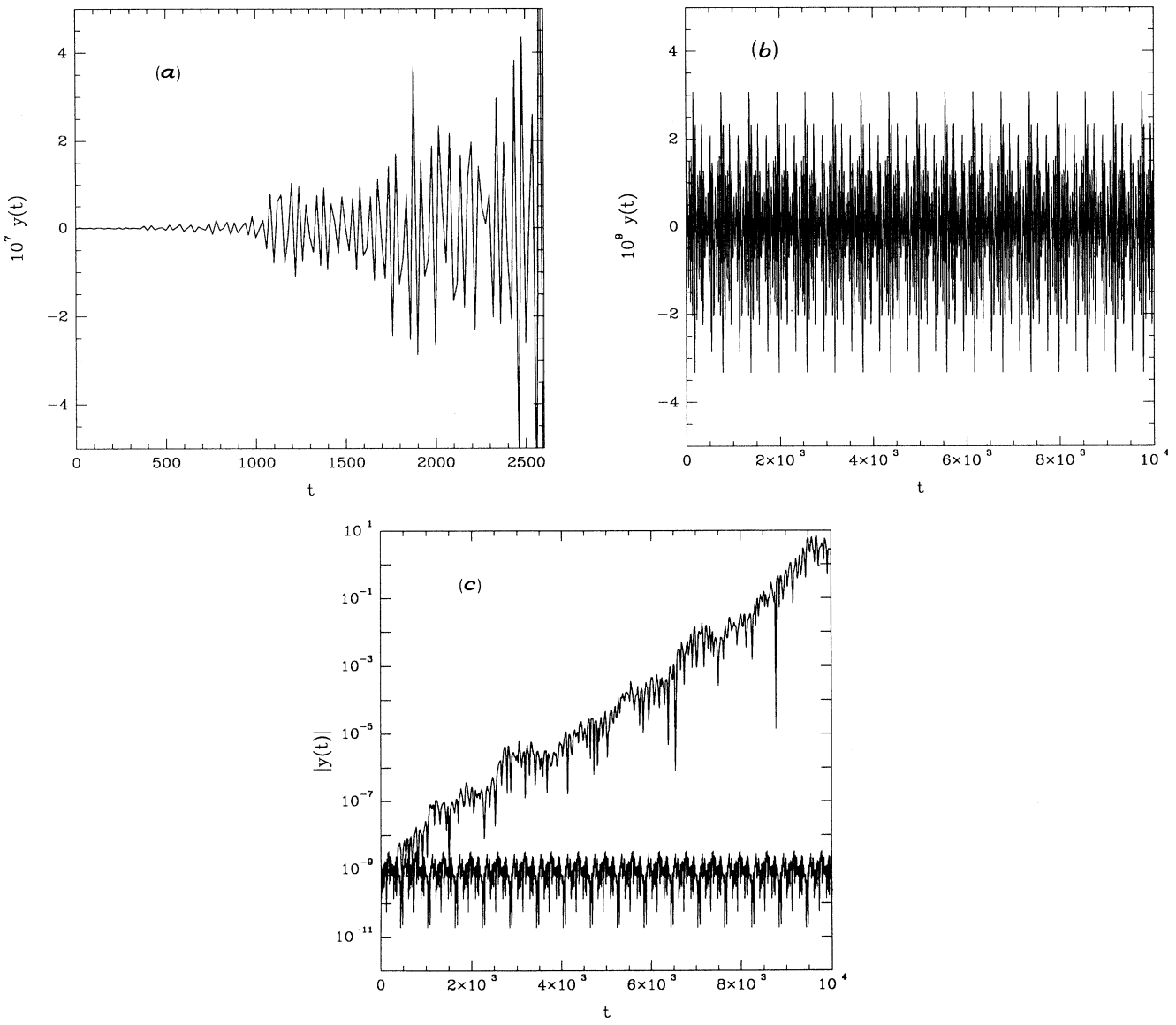


FIG. 4. Numerical solutions of Hill's equation (63) for the FPU model at $n = 128$. (a) $E = 2570$, initial conditions: $y(0) = 10^{-9}$, $\dot{y}(0) = 0$; (b) same energy and initial conditions of the previous case but $\mu = 0$ (harmonic limit); (c) the absolute values $|y(t)|$ of the solutions (a) and (b) are plotted in log-lin scale: an exponential growth is clearly evident in the chaotic case while bounded oscillations are evident in the integrable case.

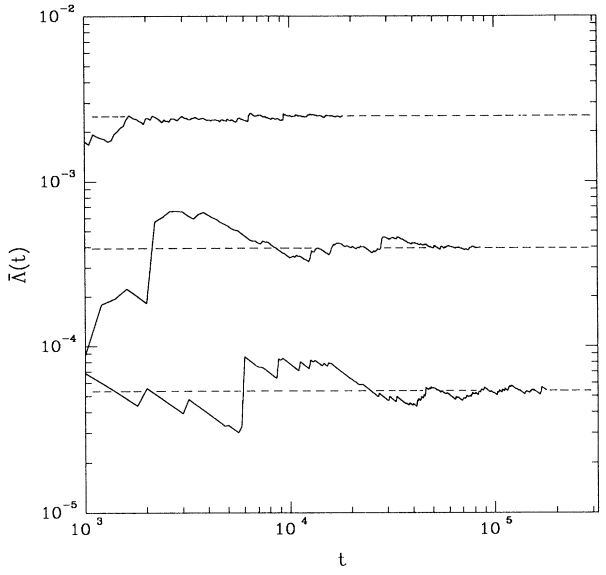


FIG. 5. Examples of the time behavior of the running averages $\bar{\Lambda}(t)$ for FPU at (from top to bottom): $E=6989$, 447, and 47.

worth noticing that $\Lambda(\varepsilon)$ suggests a geometrical origin of the transition between weak and strong chaos while this is not the case of $\lambda_1(\varepsilon)$.

Moreover, Λ has also another interesting feature: it is sensitive even to a very weak instability of the trajectories. This is well illustrated through an example provided by the FPU model. In Fig. 7(a) the relaxation pattern is reported of the largest numerical Lyapunov exponent $\lambda_1(t)$ obtained at $n=128$, $\ln\varepsilon=-5$ ($E=0.862$) and with random initial conditions at equipartition. To until $t\simeq 2\times 10^6$ it is found $\lambda_1(t)\sim t^{\alpha}/t$ with $\alpha\ll 1$, therefore, in case of insufficiently long integration time, such a behavior—close to $\lambda_1(t)\sim t^{-1}$—could erroneously suggest the existence of a regular region in phase space at low energy. On the other hand, the presence of the slow diffusion term t^{α} gives an average divergence rate of nearby trajectories following a stretched exponential: $\|\xi(t)\|=\|\xi(0)\|\exp[(t/\tau)^{\alpha}]$, with τ a constant. Such a law suggests an intermittent diffusion in phase space due to the existence of random trappings of the trajectories with a hierarchical distribution of trapping times (see the first of Ref. [14]); finally, at $t>2\times 10^6$, a "big stochastic sea" is reached and λ_1 begins to fluctuate around some nonvanishing value.

Let us now look at Fig. 7(b), where the number of times $N^{(-)}(t)$ that $\Re(\dot{\mathbf{q}})<0$ is plotted *versus* time for the same parameters and initial condition of Fig. 7(a). One immediately realizes that (i) even with a much shorter integration time it is unambiguously evident that chaos is present, and (ii) the above-mentioned intermittent diffusion shows up in a very simple way; in fact, the staircase structure of $N^{(-)}(t)$ is very pronounced: long flat pieces corresponding to finite amplitude oscillations of nearby trajectories alternating with jumps corresponding to regions of exponential instability of the dynamics. In the language of resonances, it is reasonable to think that each *plateau* appears when diffusion occurs *along* resonances, and that the jumps correspond to the *crossing* of resonances.

Even though the integration time of Fig. 7(b) is far from being sufficient to get a stabilized average of $\bar{\Lambda}(t)$, the information about chaoticity of the dynamics is neat. In other words, $\Re(\dot{\mathbf{q}})$ seems to convey more information than the largest eigenvalue of $\mathcal{H}(V)$, i.e., than numerical Lyapunov exponent λ_1. In the light of Eq. (53) this statement might be reasonable, in fact $\Re(\dot{\mathbf{q}})$ contains other in-

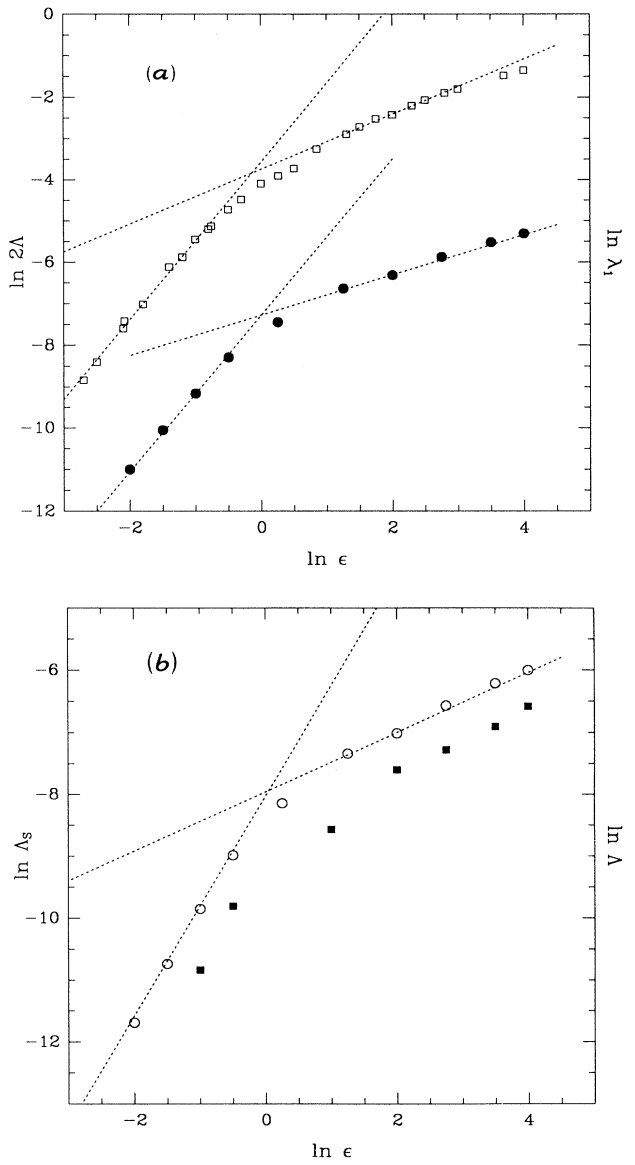


FIG. 6. FPU model. (a) Synopsis of $\Lambda(\varepsilon)$ (full circles) and $\lambda_1(\varepsilon)$ (open squares) for FPU. Their crossovers are in very good agreement and define ε_c of the SST. (b) Synopsis of $\Lambda_s(\varepsilon)$ (full squares) and $\Lambda(\varepsilon)$ (open circles), qualitatively showing that also $\Lambda_s(\varepsilon)$ is sensitive to the SST.

gredients besides $\mathcal{H}(V)$.

In a first rough approximation, the results reported in Figs. 2–4 and 7(b) suggest a picture of the dynamics where chaoticity is due to the existence of "scattering" regions on M or TM. These "scattering" regions are those subsets of M where $R<0$ or $\Re(\dot{\mathbf{q}})<0$, and where initially close trajectories undergo an exponential increase of the norm of their separation. These scattering regions seem rather small in size and at large scale they appear as rather uniformly distributed. We mention that the fluctuations of the local density of these scattering regions could provide a geometrical picture for the intermittent fluctuations in Hamiltonian chaotic dynamics that are phenomenologically described by generalized Lyapunov exponents [34].

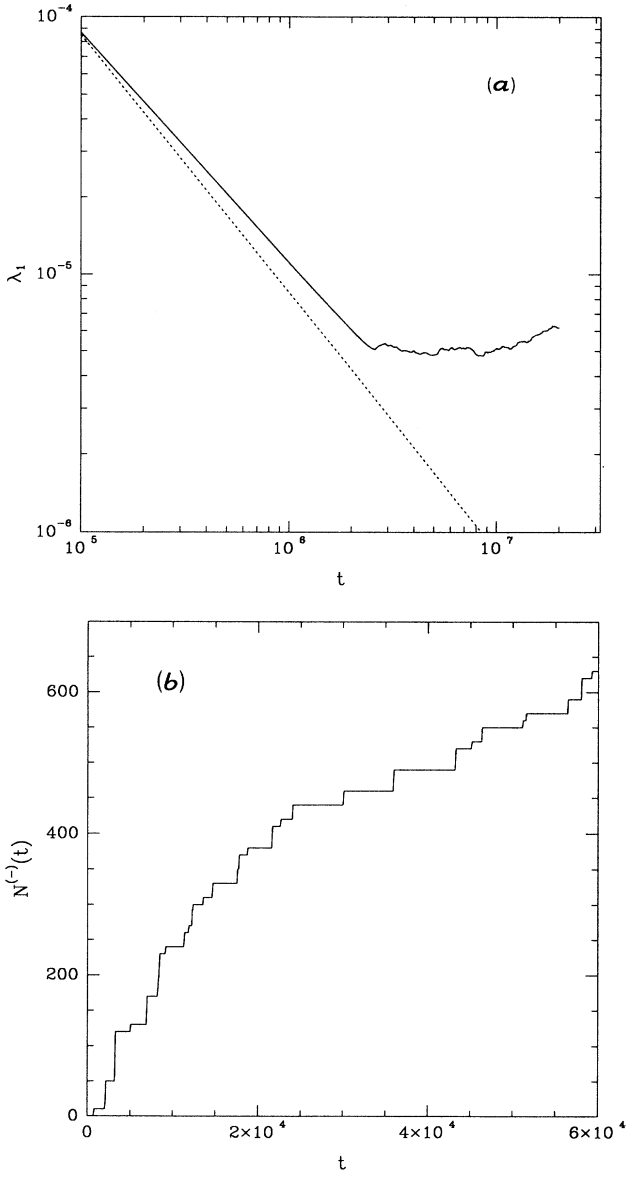


FIG. 7. FPU model. (a) Relaxation of the largest Lyapunov exponent at $E=0.862$ ($\ln\varepsilon=-5$). The reference dotted line shows a t^{-1} behavior (typical of ordered motion). Random initial conditions. (b) At the same energy and with the same random initial conditions, the number of times that Ricci curvature is found negative is reported as a function of time.

Let us remind that such a picture is derived within the approximations performed on the Jacobi–Levi-Civita equation in the preceding section. It is certainly possible that the largest eigenvalue of the operator defined in Eq. (42) happens to be negative in regions of larger measure than those where $R<0$ or $\Re(\dot{\mathbf{q}})<0$. Moreover, also the n equations $(d^2\xi^i/ds^2)+\lambda_{(i)}\xi^i=0$, obtained by diagonalizing Eq. (43), are most probably subject to parametric modulation, i.e., $\lambda_{(i)}=\lambda_{(i)}(t)$, because of the already mentioned "bumpiness" of M. Hence, even when all the $\lambda_{(i)}$ are positive, still an unstable dynamics could be obtained, provided that the conditions for parametric resonance are fulfilled.

It is worth noticing that, in the average, both the scalar and Ricci curvatures are positive. These quantities enter into the definition of GCI after the already reminded approximations. It was the scalar curvature [Eq. (54)] that drew Krylov's attention in the cited work of Ref. [16]. Krylov *heuristically* inferred the mixing behavior of a Hamiltonian system from the condition that $R<0$ almost everywhere. However, this is an oversimplification of the problem: we have already seen that another active source of chaos is provided by the bumpiness of the ambient manifold through parametric resonance, regardless of the sign of the curvature.

Moreover, it is not worth paying too much attention to scalar curvature in the study of chaos. In fact, let us consider the example of $(M\times\mathbb{R}^2,g_E)$ as ambient manifold for the dynamics, it *always* holds true that $R=0$, whereas this is not the case of $\Re(\dot{\mathbf{q}})$. Another drawback of the scalar curvature shows up in numerical simulations at large n (up to $n=10\,000$) where an "abnormal" increase is found of the number of regions where $R<0$; this happens even at small ε when chaos—detected by other means—is weak. At variance, $\Re(\dot{\mathbf{q}})$ is exempt form this problem; this is also because $\Re(\dot{\mathbf{q}})$ contains more information than R: it is obtained by one instead of two averagings and involves $n^2/2$ instead of n terms, besides, it is built up of a nontrivial saturation of the off-diagonal components of R_{ik} with the velocities. However, here the situation is reversed: some preliminary tests at increasing n and constant ε give a corresponding increase of the mean free path between two successive encounters of scatterers [regions of $\Re(\dot{\mathbf{q}})<0$]. A systematic investigation about the n dependence of our results is beyond the aims of the present paper: the previous remarks aim at warning again about the partial loss of information which is produced by any averaging of Eq. (35). *A priori* only the eigenvalues of the operator Q in Eq. (43) and their time averages could keep all the necessary information to provide GCI's free of the above-mentioned problems. Nevertheless, in what follows a simple way is shown to recover some of the lost information.

We wish to find some quantity that, being easily computable, can detect the existence of a SST, can discriminate among integrable and nonintegrable systems and, obviously, has a geometrical meaning. In other words, we seek some major change in the geometry of (M,g_J)

that could be interpreted as a geometrical mark of the SST.

The simplest choices at hand are the following:

$$\hat{R}=\frac{1}{n(n-1)T}\int_0^T dtR[q^a(t)] \tag{79}$$

and

$$\hat{G}=\frac{1}{nT}\int_0^T dtR_{ik}[q^a(t)]\dot{q}^i(t)\dot{q}^k(t)\ , \tag{80}$$

i.e., $\hat{R}$ and $\hat{G}$, respectively, measure the overall scalar and Ricci curvatures by averaging along a trajectory and regardless of the sign of the curvature. The striking result is that also $\hat{R}(\varepsilon)$ and $\hat{G}(\varepsilon)$ neatly detect the SST.

In Fig. 8 $\hat{R}(\varepsilon)$ is reported in the case of the FPU model. The convergence in time of $\hat{R}$ is very fast and with small fluctuations. For instance, at $E=4239$ ($\ln\varepsilon=3.5$) the convergence time is ~ 300, and at $E=36.6$ ($\ln\varepsilon=-1.25$) the convergence time is ~ 4000. Both $\hat{R}(\varepsilon)$ and $\hat{G}(\varepsilon)$ are computed together with Λ (which has a slower relaxation rate), thus the averaging time T is so long that the values of $\hat{R}(\varepsilon)$ and $\hat{G}(\varepsilon)$ are obtained with high precision. At $\varepsilon<\varepsilon_c$, it is found that $\hat{R}(\varepsilon)\sim\varepsilon^{-2}$, while at $\varepsilon>\varepsilon_c$ a sharp transition shows up to a less steep power law; here ε_c is in perfect agreement with the critical energy density of the SST mentioned above. The broken line on the background is obtained by setting $\mu=0$. Also at small ε the numbers $\hat{R}_{\mu=0}(\varepsilon)$ and $\hat{R}(\varepsilon)$ are different, but this difference cannot be appreciated in Fig. 8 for graphical reasons. Notice that $\hat{R}$ is a decreasing function of ε, while the instability exponent $\chi=2RW^2/n(n-1)$ is an increasing function of ε.

A similar result, though more pronounced, is obtained through $\hat{G}(\varepsilon)$ and is reported in Fig. 9 for the FPU model. Also in this case, $\hat{G}$ is quickly convergent with small fluctuations. The convergence rates and fluctuation amplitudes are those of $\hat{R}$. The critical value of the threshold, defined by the crossing of the two asymptotes, coincides with the preceding value obtained either with $\lambda_1(\varepsilon)$ or with $\Lambda(\varepsilon)$ or with $\hat{R}(\varepsilon)$.

By setting $\mu=0$, $\hat{G}(\varepsilon)$ is found perfectly constant, as shown in Fig. 9. This provides a clear-cut criterion to recognize regular and chaotic dynamics. The asymptotic convergence of $\hat{G}(\varepsilon)$ to $\hat{G}_{\mu=0}(\varepsilon)$ at small ε is indicative of weak chaoticity, not of absence of chaos; as already discussed, chaoticity is always present.

In Fig. 10 $\hat{G}(\varepsilon)$ and $\hat{G}_{\mu=0}(\varepsilon)$ are reported in the case of φ^4 model. A qualitatively similar result to that of the FPU case is obtained. With the chosen values of m^2 and μ, ε_c is found smaller than in the FPU case, in agreement with what is given by $\lambda_1(\varepsilon)$ [14]. In Ref. [14], the esti-

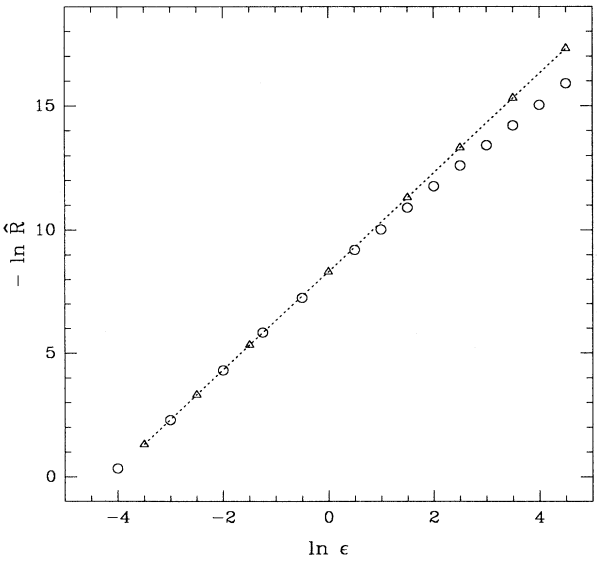


FIG. 8. FPU model. Time averages of the scalar curvature [Eq. (79)] $-\ln\hat{R}$, are reported vs $\ln\varepsilon$. The triangles, joined by the dotted line, refer to $\mu=0$ (harmonic limit). The open circles refer to the chaotic case ($\mu\neq 0$) and show the existence of the SST at the good ε.

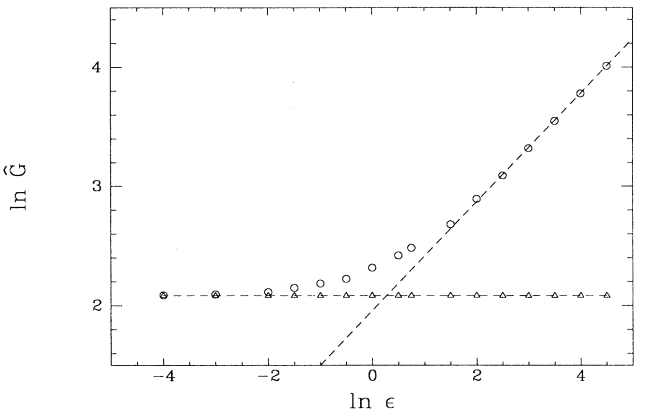


FIG. 9. FPU model. Time averages of the Ricci curvature [Eq. (80)] $\ln\hat{G}$ are reported vs $\ln\varepsilon$ (open circles). The triangles refer to the harmonic limit ($\mu=0$). The existence of the SST is here neatly evident.

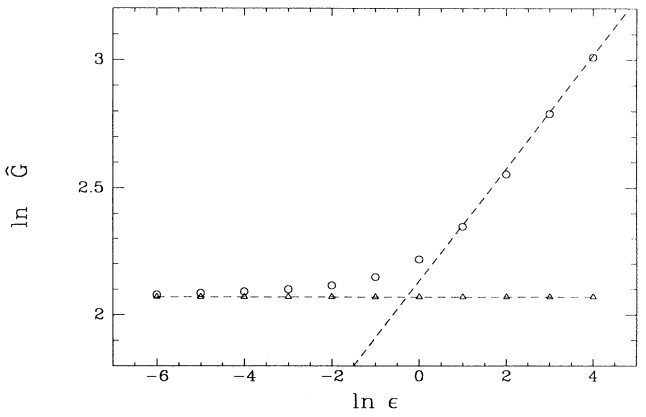


FIG. 10. φ^4 model. Time averages of the Ricci curvature [Eq. (80)] $\ln\hat{G}$ are reported vs $\ln\varepsilon$ (open circles). The triangles refer to the harmonic limit ($\mu=0$). Also in this case the SST is clearly detected.

mate of ε_c, in the case of the φ^4 model, is lower than the value given by Fig. 10, but the data of Ref. [14] are absolutely compatible with the sharper result of Fig. 10. Another qualitative agreement between $\hat{G}(\varepsilon)$ and $\lambda_1(\varepsilon)$ is found by comparing the results for FPU and φ^4 models. In fact, in the φ^4 case, the transition occurs between $\lambda_1 \sim \varepsilon^{3/4}$ and $\lambda_2 \sim \varepsilon^{2/3}$, i.e., it is less sharp than in the FPU case where $\lambda_1(\varepsilon)$ passes from $\lambda_1 \sim \varepsilon^2$ to $\lambda_1 \sim \varepsilon^{2/3}$; similarly, $\hat{G}(\varepsilon)$ passes from ε^0, in both cases, to $\varepsilon^{1/5}$ (φ^4 model) and to $\varepsilon^{1/2}$ (FPU model).

In Fig. 11 $\hat{G}(\varepsilon)$ is reported for the Toda lattice. As in the case of harmonic oscillators $\hat{G}(\varepsilon)$ turns out perfectly constant, so confirming that a simple but nontrivial possibility exists to recover some of the information lost in the averaging of Eq. (35). The ε dependence of the geometrical quantities chosen provides this retrieval of information.

Even though we have only numerical evidence for the following statement, nonetheless we can tentatively suggest that if

$$\frac{d\hat{G}(\varepsilon)}{d\varepsilon}=0 , \tag{81}$$

then we are dealing with an integrable system. An equivalent statement holds for $d[\varepsilon^2\hat{R}(\varepsilon)]/d\varepsilon=0$.

The neat results obtained with $\hat{R}$ and $\hat{G}$ seem to mirror some deeper relationship between geometrical and dynamical properties of Hamiltonian flows. In the next section we shall briefly comment on this point.

Finally, let us come to the comparison between dynamic and static averages. Both $\hat{R}$ and $\hat{G}$ are well suited for static averaging, which is not the case, for example, of Λ, which has a much slower convergence rate. Static averaging is defined as in Eq. (65) by

$$\langle R\rangle=\Omega^{-1}\int_{\Sigma_E} d\sigma_E R(q^a)/n(n-1)$$
$$=\Omega^{-1}\int\prod_{i=1}^{n} dq_i d\dot{q}_i \delta(H'(q^a,\dot{q}^a)-E)R(q^a)/n(n-1) \tag{82}$$

and

$$\langle G\rangle=\Omega^{-1}\int_{\Sigma_E} d\sigma_E R_{hk}(q^a)\dot{q}^h\dot{q}^k/n$$
$$=\Omega^{-1}\int\prod_{i=1}^{n} dq_i d\dot{q}_i \delta(H'(q^a,\dot{q}^a)-E)$$
$$\times R_{hk}(q^a)\dot{q}^h\dot{q}^k/n . \tag{83}$$

Numerical computation of these surface integrals has been performed by a standard Monte Carlo technique adapted to the constraint: $\delta(H'(q^a,\dot{q}^a)-E)$ [35]. To make the results more reliable, two different algorithms have been used. The first one [35] employs a Maxwellian demon that exchanges energy with the system under investigation. The total energy $E=E_d+E_s$ is kept constant, whereas the demon energy E_d and the system energy E_s vary. Only positive values of E_d are allowed. Once the equilibrium is established, if the demon is "very small," then the most likely value of E_s is very close to E. In analogy with a conventional Monte Carlo–Metropolis simulation, each variable q_i and $\dot{q}_i$ is systematically varied and the energy of the new configuration is compared with the preceding one. If energy is gained then the system moves to the new configuration, and the energy is given to the demon. If energy is required by the update, then the system moves only if the demon has enough energy, and the energy needed is taken from the demon. If the demon has not enough energy, no change occurs.

The second algorithm consists of a Gaussian representation of the δ function, i.e., $\exp[-(H-E)^2/\Gamma]$, with Γ a parameter to be conveniently chosen. A new configuration is accepted if its probability of occurrence, computed by this Gaussian, is greater than a random number extracted in the interval [0,1], otherwise, it is rejected.

In both cases a random initial configuration $\{\mathbf{q}^{(0)},\dot{\mathbf{q}}^{(0)}\}$ at equipartition, like in Eq. (76), is dynamically evolved up to a time $t=1000$ that ensures virialization. This is done in order to initialize the Monte Carlo simulation with a configuration that has both random velocities and positions. The updating procedure is sequential (some tests have been also made by choosing at random the coordinates), and follows the simple scheme

$$\dot{q}_i^{(k+1)}=\dot{q}_i^{(k)}+\alpha\sqrt{\varepsilon}w_k, \quad q_i^{(k+1)}=q_i^{(k)}+\alpha\sqrt{\varepsilon}w_k' , \tag{84}$$

where α is a numerical factor needing an empirical optimization, w_k and w_k' are Gaussian random numbers with zero mean and unit variance; and the factor $\sqrt{\varepsilon}$ is adopted as an empirical scaling with the energy density: this turned out useful to keep α in a small range of working values.

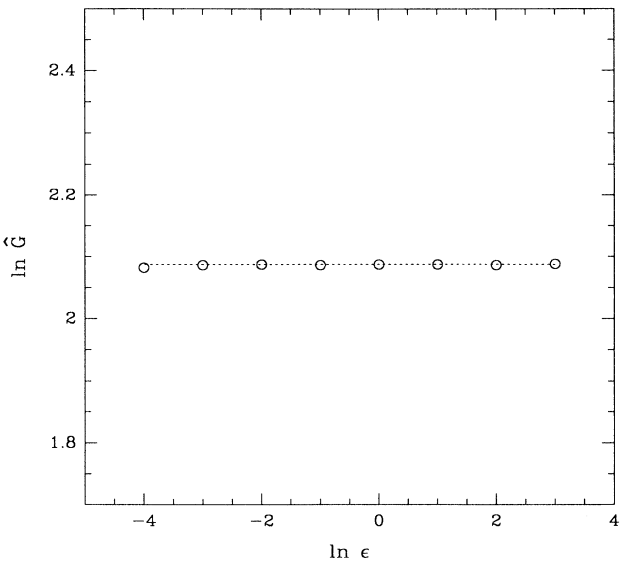


FIG. 11. Toda lattice. Integrable case. Time averages of the Ricci curvature [Eq. (80)] $\ln\hat{G}$ are reported vs $\ln\varepsilon$ (open circles). The result is the same as in the case of harmonic limits of FPU and φ^4 models.

The above-described algorithms constrain the random walk given by Eq. (84) only *near* a constant-energy surface. Therefore, by a suitable correction of the coordinates, all the accepted new configurations are projected on the initial energy surface Σ_E: $H'(q^a,\dot{q}^a)=E$. The procedure goes as follows. Define: $\mathbf{x}=\{x_1,\ldots,x_{2n}\} \equiv\{q_1,\ldots,q_n\dot{q}_1,\ldots,\dot{q}_n\}$. If at the kth Monte Carlo step $\mathbf{x}^{(k)}$ is a configuration such that $\mathbf{x}^{(k)}\in\Sigma_E$, at the $(k+1)$th Monte Carlo step the new configuration is in general such that $\tilde{\mathbf{x}}^{(k+1)}\in\Sigma_{E+\Delta E}$, therefore we compute a correction $\Delta\mathbf{x}^{(k+1)}$ so that $\mathbf{x}^{(k+1)}=\tilde{\mathbf{x}}^{(k+1)} +\Delta\mathbf{x}^{(k+1)}\in\Sigma_E$ by simply projecting $\tilde{\mathbf{x}}^{(k+1)}$ on Σ_E. This is easily achieved by minimizing the Euclidean distance

$$d^2=\sum_{i=1}^{2n}(\Delta x_i^{(k+1)})^2$$

with the constraint $\mathbf{x}^{(k+1)}\in\Sigma_E$. Thus from

$$\Delta E=\sum_{i=1}^{n}\left[\frac{\partial E}{\partial q_i}\Delta q_i+\frac{\partial E}{\partial p_i}\Delta p_i\right]=\sum_{i=1}^{2n}\frac{\partial H}{\partial x_i}\Delta x_i\ , \quad (85)$$

we write

$$\frac{\partial}{\partial(\Delta x_i)}\left[\sum_{i=1}^{2n}(\Delta x_i)^2+\lambda\left[\sum_{i=1}^{2n}\frac{\partial H}{\partial x_i}\Delta x_i-\Delta E\right]\right]=0\ ; \quad (86)$$

this gives $\Delta x_i=-\frac{1}{2}\lambda(\partial H/\partial x_i)$, so that introducing it into Eq. (85), we get

$$\lambda=-\frac{2\Delta E}{\Sigma_i\left[\frac{\partial H}{\partial x_i}\right]^2} \quad (87)$$

and finally

$$\Delta x_i^{(k+1)}=\left[\frac{\Delta E\left[\frac{\partial H}{\partial x_i}\right]}{\sum_{i=1}^{2n}\left[\frac{\partial H}{\partial x_i}\right]^2}\right]_{\tilde{\mathbf{x}}^{(k+1)}}\ , \quad (88)$$

which are the corrections needed to project $\tilde{\mathbf{x}}^{(k+1)}$ on Σ_E.

This correction scheme is added to both of the mentioned Monte Carlo (MC) algorithms and turns out to be a simple and efficient way to generate a random walk on any constant-energy surface Σ_E. This random-walk samples Σ_E on a set of points that is used to compute $\langle R\rangle$ and $\langle G\rangle$. In order to measure how the structure of Σ_E changes with energy, several computations are performed at different values of E. Notice that $\langle G\rangle$ measures the total Ricci curvature of (M,g_J) but it has not an equivalent meaning for Σ_E: $\Re(\dot{\mathbf{q}})$ is only a function defined on Σ_E.

Let us now give some details about the numerical parameters used. Unless explicitly stated, things refer to both algorithms.

If the parameter α is chosen too small, the diffusion process on Σ_E is very slow and so is the convergence rate of the averages. On the other hand, if α is large, the deviations from Σ_E are too wide; consequently, the acceptance rate of new configurations drops down and the projection technique becomes less precise. It has been empirically found that the good values fall in the interval [0.05,0.5]; the smaller ones are used at higher energies. In the case of the demon algorithm, $\alpha=0.1$ is optimal at all the energies considered. The average acceptance rate of the MC trials ranges from about 50% to about 35% (see below). In the case of the Gaussian approximation of the δ function, the parameter Γ is chosen according to the empirical rule $\Gamma=\varepsilon\alpha^2\beta$, where β is varied in the interval [1,100]; again larger β are used at higher energies. Unfortunately the good choices of the parameters (see below) give average acceptance rates of the MC trials of abut 4%, which makes this method by far less efficient.

In both cases, energy is conserved by the projection procedure, with a maximum relative error of 10^{-3}. After completion of one trial per particle on the whole set of coordinates, each updated configuration is projected on Σ_E and so corrected, enters the set of configurations used to compute the averages.

A simple criterion has been adopted to assess the reliability of the MC averages: in any MC run also the average potential energy $\langle V\rangle_{\rm MC}$ is computed, and then it is compared with the analytically known value $\langle V\rangle_{\rm ens}$ [36]. In Ref. [36] $\langle V\rangle_{\rm ens}$ is computed in the canonical ensemble for the FPU model; at $n=128$ the canonical mean is a good approximation of the microcanonical one. The average kinetic energy per particle provides the temperature. *A priori* different observables may have very different convergence times (both dynamical and MC), depending on their degree of smoothness on Σ_E [36]. For instance, a quantity that depends on fluctuations like the specific heat needs a much denser sampling than a quantity involving only simple and regular functions of the

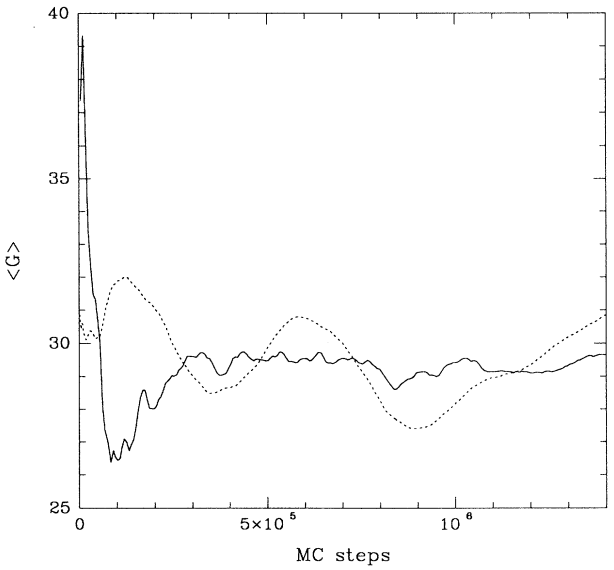


FIG. 12. FPU model. Relaxation of the Monte Carlo averages [Eq. (83)] of the Ricci curvature $\langle G\rangle$ at $E=3301$. Comparison of the two methods used: demon algorithm (full line) with $\alpha=0.1$, and Gaussian representation of the δ function (dotted line) with $\alpha=0.1$, $\beta=1$.

coordinates. In particular, G and V have a qualitatively comparable degree of regularity of their functional dependences upon the coordinates [i.e., algebraic functions of $(q_{k+1}-q_k)$]. So a good convergence of $\langle V\rangle_{\rm MC}$ can be taken as the mark of a sufficient uniformity in the sampling of Σ_E to make $\langle G\rangle_{\rm MC}$ reliable.

A comparative example of the results obtained with the two methods is given in Fig. 12. The agreement is fairly good, though the demon algorithm is faster, yields smaller fluctuations of the running average and makes the convergence of $\langle V\rangle_{\rm MC}$ to $\langle V\rangle_{\rm ens}$ quicker. The results obtained with the approximate δ function algorithm are mainly used to further confirm the reliability of the output of the demon method.

To get a hold of the convergence rate and of the precision of the final results, in Fig. 13 three different examples are given of $\langle G\rangle$ versus the number of MC steps. They refer to $E=3301$, 737, and 22, that is, $\ln\varepsilon=3.25$, 1.75, and -1.75, respectively. The total number of configurations involved in the averaging are around 28 000, 34 000, and 39 000, respectively; in all these cases 10^7 MC trials have been done, so from the abscissas of the end points of the three curves one can immediately figure out the acceptance rates.

The convergence of $\langle R\rangle$ is faster: already with few thousands of configurations a good approximation of the final result is obtained. Instead of giving another picture, where the graphical squeezing would hinder the appreciation of the excellent agreement between dynamic and static averages, let us give a few examples. At $E=6989$, 348, and 47, that is, $\ln\varepsilon=4$, 1, and -1, respectively, it is found that $\hat{R}=0.2983\times10^{-6}$, 0.4431×10^{-4}, and 0.1922×10^{-2} for time averages and $\langle R\rangle=0.297\times10^{-6}$, 0.445×10^{-4}, and 0.191×10^{-2} for MC averages (only stabilized digits are given).

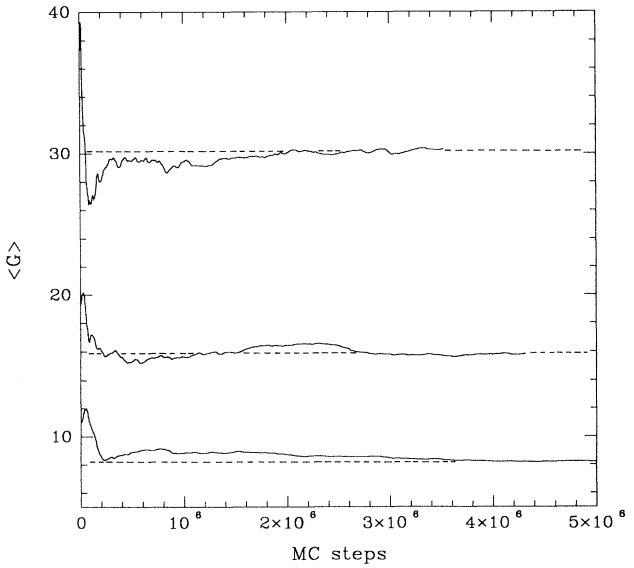


FIG. 13. FPU model. Relaxation of the Monte Carlo averages $\langle G\rangle$ obtained with the demon algorithm at $E=3301$, 737, and 22 (from top to bottom).

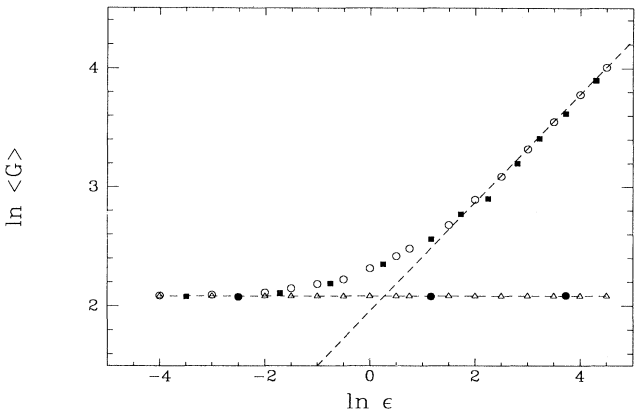


FIG. 14. FPU model. Comparison among dynamical and static (time and Monte Carlo) averages of the Ricci curvature vs ε. Full squares and circles show the Monte Carlo averages in the chaotic and harmonic cases, respectively. Open circles and triangles represent time averages.

Finally, in Figs. 14 and 15 $\langle G\rangle$ versus ε is reported for FPU and φ^4 models, respectively; a comparison is also made with time averages of G. The striking result is that dynamic and static averages are in excellent agreement. Certainly this result is not obvious. It suggests that the *total* Ricci curvature [i.e., its integral on (M,g_J)] is related to some deeper geometrical property which is responsible for the regular or chaotic behavior of the geodesics of M.

Both $\hat{G}$ and $\langle G\rangle$ are computed on M with the condition that positions and momenta must belong to a given section TM_E of TM, so the transitional behavior revealed by $\hat{G}(\varepsilon)$ and $\langle G\rangle(\varepsilon)$ is ascribed to a rather abrupt change of the structure of the constant-energy surfaces of TM at some critical energy, which perhaps could mean that a

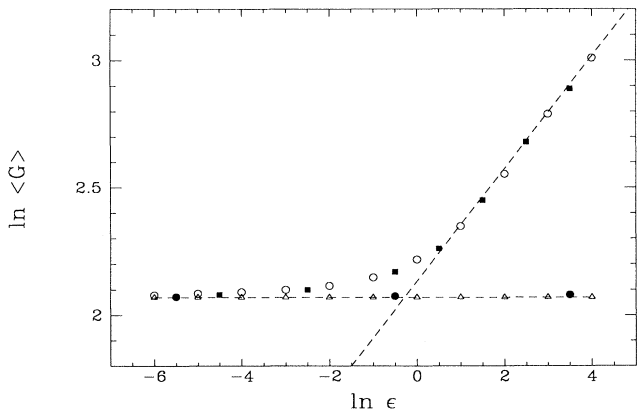


FIG. 15. φ^4 model. Comparison among dynamical and static (time and Monte Carlo) averages of the Ricci curvature vs ε. Full squares and circles show the Monte Carlo averages in chaotic and harmonic cases, respectively. Open circles and triangles represent time averages.

singular energy foliation of TM exists.

This fact deserves a lot of theoretical work about one-parameter deformations of Riemannian metrics and the corresponding integrability of the resulting geodesic motions, whereas, for physical applications, we have already found—at least in principle—a method to compute ε_c of the SST by a *static* algorithm: by averaging G on the surfaces $\{\Sigma_E\}$, one has to find out at what energy $\langle G\rangle(\varepsilon)$ displays its transition. This could be achieved in several empirical ways. For instance, if the high-energy behavior of $\langle G\rangle(\varepsilon)$ can be approximated by some function $\langle G_{\rm HE}\rangle(\varepsilon)$, and if $\langle G_0\rangle(\varepsilon)$ refers to the case $\mu=0$, then ε_c can be obtained by solving $\langle G_{\rm HE}\rangle(\varepsilon)=\langle G_0\rangle(\varepsilon)$. Another possibility consists in finding the ε value at which $(d^3\langle G\rangle d\varepsilon^3)=0$. In fact, the function $(d\langle G\rangle/d\varepsilon)$ suddenly increases in the transition region, so its inflection point can be used to define ε_c.

IV. CONCLUSIONS

The central point of the present paper concerns the practical use of elementary tools of Riemannian differential geometry to describe the origin of Hamiltonian chaos. Riemannian geometry enters only in the study of geodesic flows within the framework of ergodic theory, but these are abstract flows that do not help very much in understanding the chaotic properties of Hamiltonian flows of physics. Here we have shown that the geometrical approach is actually useful for these flows of physical interest. In particular, it is possible to go beyond the usual picture of homoclinic intersections to explain the origin of Hamiltonian chaos. This is done in a quite natural way. In fact everything stems out of the variational formulation of Newton's law of dynamics. Then a classical tool of Riemannian geometry is used to investigate the stability properties of motion: the Jacobi–Levi-Civita equation for the geodesic spread. In particular, written in the case of Eisenhart's metric on $M\times\mathbb{R}^2$, this yields the equation for the tangent dynamics that is commonly used to compute numerical Lyapunov exponents. Hence a new interpretation is given to the meaning of the results of the standard algorithm [30]. It turns out that there is no need to invoke Oseledets's theorem: numerical Lyapunov exponents are not true Lyapunov exponents, nevertheless, they have their own well-defined meaning.

The approach adopted in this paper is naturally *nonperturbative*. In fact, no references to nearly integrable systems are made, and all the equations or quantities used are well defined at any energy. Moreover, there is no need in action-angle coordinates because natural coordinates and velocities are used. This is an important advantage for the overwhelming majority of physical systems for which the computation of action-angle coordinates is long and tedious already at low perturbative order and at small n. In any case, action-angle coordinates belong to the machinery of CPT and therefore these are powerful within the validity limits of CPT.

Our investigation of Hamiltonian chaos by Riemannian geometry has been motivated by the need of finding an explanation for the existence of the SST. As already discussed, this cannot be understood within the framework of CPT. In the present paper we have successfully rephrased the problem in geometrical terms. Of course, this is not yet a deep mathematical explanation, but this is much more than pure phenomenology. In fact, it opens a new perspective of research for high-dimensional Hamiltonian flows by revealing that some nontrivial structure related with the energy foliation of TM is responsible for the existence of a critical energy at which a major geometrical change occurs. Correspondingly, the topology of the phase-space paths is deeply affected. This is at the origin of weak and strong chaos, and of slow and fast diffusion in phase space. A more stringent mathematical definition of "geometrical change" is necessary. One could wonder whether some topological change—defined through the usual homological approach—might be related with the existence of the SST. Let us first consider (M,g_J). We can expect two opposite situations: chaos will appear on manifolds of complex topology (loosely speaking some n-dimensional analog of high genus); ordered motion will require trivial homology. The former could be a sufficient topological condition to make chaos, the latter could be a necessary condition for integrability. From general theorems [31], we know that if $\Re(\dot{\mathbf{q}})>0$ for each $\mathbf{q}\in M$, and if M is compact and conformally flat [which is the case of (M,g_J)], then M is a homology sphere. This seems in agreement with the numerical results obtained for integrable systems. However, even smooth deformations of the metric of a manifold—whose geodesics are stable—can yield dramatic effects on their stability because geodesics are not preserved by diffeomorphisms of the underlying manifold. This fact and the above reported results about parametric resonance due to the "bumpiness" of (M,g_J) suggest that we can have chaos also on a homologically trivial manifold. In other words, trivial homology is certainly not a sufficient condition for integrability, and if we think of geodesic flow of the usual torus immersed in $\mathbb{R}^3$, we see that trivial homology is also not necessary for integrability. Regular and chaotic motions depend on the differentiable structure defined by the metric of the manifold. Thus geometrical rather than topological properties are relevant, and so the necessity appears for a deeper characterization of geometrical properties and their changes. The use of differential invariants [27] and of the corresponding generalized cohomological complexes [37] and characteristic classes are very promising mathematical concepts and methods of tackling the above-discussed problem.

Let us now briefly discuss the constant-energy surfaces of TM. Also the tangent bundle TM of configuration space can be equipped with a Riemannian metric. There are many possibilities [38] of lifting the metric of the base manifold M to TM. However, among the others, the so-called Sasaki diagonal lift g_S [39] is the most natural. The manifold (TM,g_S) is a Riemannian manifold. The geodesics of (TM_E,g_S)—with suitable restrictions—are the natural motions on a constant-energy surface of the dual of phase space. In this case it is an open question whether some topological change—in the above-discussed sense—occurs at the SST.

The second of the two main concerns of the present

work, as stated in the Introduction, is about finding an algorithm to compute ε_c of the SST. By static averaging of geometrical quantities, we find that something abruptly changes in the geometry of the constant-energy surfaces of TM, and this happens at the same energy of the SST. The agreement between dynamic and static averages allows a static computation (i.e., of the statistical-mechanical kind) of the SST. Simplified criteria are suggested to compute the critical energy density ε_c. Obviously, this point also deserves further work.

The results reported in this paper suggest that a Riemannian description of Hamiltonian chaos is worthwhile. Certainly the approach proposed here opens many new problems. However, it has the great advantage of bridging with a largely developed topic of mathematics that could supply Hamiltonian dynamics with new methods and results.

In Sec. II it is only mentioned that differential geometry of Finsler spaces could be of interest to Hamiltonian dynamics. This is a subject that deserves further attention and investigation. Just to give examples of physical systems requiring a Finslerian approach, notice that conservative and velocity-dependent forces, like the Lorentz force, are naturally described by Finsler spaces. This is the case of systems subject to the action of electrodynamical interactions that naturally arise, for example, in plasma physics problems [40]. Other interesting cases (also in view of applications) are given by the problem of diffusion of passive scalars in nonstationary Eulerian velocity fields of fluids [41] or, equivalently, of diffusion of charged particles in fusion plasmas in the guiding centers approximation [42]. In both cases the equations of motion are $\dot{x} = -\partial_y U(x,y,t)$ and $\dot{y} = \partial_x U(x,y,t)$, where $U(x,y,t)$ does not have a Newtonian structure; that is, the kinetic part is lacking and so a Riemannian approach is hindered, whereas the function $\Lambda = \frac{1}{2}(x\dot{y} - \dot{x}y) - U(x,y,q^{n+1})\dot{q}^{n+1}$ provides a metrical function [Eq. (14)] of the Finslerian kind.

To pursue the discussion of future developments, let us mention that there is also another reason to study the above-mentioned lift of (M,g_J) to (TM,g_S): to look for new methods of describing phase-space diffusion. In fact, in order to write down a chaotic diffusion equation, it is necessary to study how a given volume of initial conditions is "geodesically advected" on TM_E. The study of chaotic diffusion in Hamiltonian systems has many important applications: the estimate of relaxation times of nonequilibrium initial states in statistical-mechanical systems, understanding the diffusion mechanisms of passive scalars in fluids or of charged particles in fusion plasmas, just to give few examples of interest to theoretical and applied physics.

ACKNOWLEDGMENTS

The author wishes to thank D. Alekseevsky, V. I. Arnold, M. Rasetti, Ya. G. Sinai, and G. Vezzosi for discussions and comments. A particularly warm acknowledgment is addressed to A. M. Vinogradov for his continuous interest in the present work and for many enlightening discussions. The overall CPU time needed in this work amounts to 85 hours on a Cray Y-MP computer. These have been financially supported by INFN Sezione di Firenze, CNR, and ENEA. All of them are gratefully acknowledged.

*Also at Istituto Nazionale di Fisica Nucleare, Sezione di Firenze, 50125 Firenze, Italy.

[1] H. Poincaré, *Les Méthodes Nouvelles de la Méchanique Celeste* (Blanchard, Paris, 1987).

[2] H. Poincaré, *Les Méthodes Nouvelles de la Méchanique Celeste* (Ref. [1]), Vol. 3, p. 389.

[3] V. K. Melnikov, Trans. Moscow. Math. Soc. **12**, 1 (1963). This method has been generalized at a higher dimension in P. J. Holmes and J. E. Marsden, J. Math. Phys. **23**, 669 (1982).

[4] E. Fermi, Nuovo Cimento **25**, 267 (1923); **26**, 105 (1923).

[5] This is explicitly stated for instance, in I. Prigogine, *Non-equilibrium Statistical Mechanics* (Wiley, New York, 1962), Chap. 14.

[6] E. Fermi, J. Pasta, and S. Ulam, in *Collected Papers of Enrico Fermi,* edited by E. Segré (University of Chicago, Chicago, 1965), Vol. 2, p. 978.

[7] A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **98**, 527 (1954); J. Moser, Nachr. Akad. Wiss. Goettinger Math. Phys. K. 2 **1**, 1 (1962); V. I. Arnold, Russ. Math. Surv. **18**, 9 (1963).

[8] L. Chierchia and G. Gallavotti, Nuovo Cimento B **67**, 277 (1982); G. Benettin, L. Galgani, and J. M. Strelcyn, *ibid.* **79**, 201 (1984); M. Vittot, Ph.D. thesis, Université de Provence, Marseille, 1985. A numerical confirmation can be found in M. Falcioni, U. Marini Bettolo Marconi, and A. Vulpiani, Phys. Rev. A **44**, 2263 (1991).

[9] N. N. Nekhoroshev, Funct. Anal. Appl. **5**, 338 (1971); Russ. Math. Surv. **32**, 1 (1977).

[10] G. Benettin and G. Gallavotti, J. Stat. Phys. **44**, 293 (1986); G. Benettin, L. Galgani, and A. Giorgilli, Celest. Mech. **37**, 1 (1985); P. Lochak, Phys. Lett. A **143**, 39 (1990) and (private communication).

[11] P. Bocchieri, A. Scotti, B. Bearzi, and A. Loinger, Phys. Rev. A **2**, 2013 (1970).

[12] B. Callegari, M. Carotta, C. Ferrario, G. Lo Vecchio, and L. Galgani, Nuovo Cimento B **54**, 463 (1979); R. Livi, M. Pettini, S. Ruffo, M. Sparpaglione, and A. Vulpinai, Phys. Rev. A **28**, 3544 (1983); **31**, 1039 (1985); R. Livi, M. Pettini, S. Ruffo, and A. Vulpiani, *ibid.* **31**, 2740 (1985).

[13] F. M. Izrailev and B. V. Chirikov, Dokl. Akad. Nauk SSR **166**, 57 (1966) [Sov. Phys.—Dokl. **11**, 30 (1966)].

[14] M. Pettini and M. Landolfi, Phys. Rev. A **41**, 768 (1990); M. Pettini and M. Cerruti-Sola, *ibid.* **44**, 975 (1991).

[15] P. Butera and G. Caravati, Phys. Rev. A **36**, 962 (1987).

[16] This can be found in his doctoral thesis of 1942, Leningrad Physical and Technical Institute, reprinted in N. S. Krylov, *Works on the Foundations of Statistical Physics* (Princeton University Press, Princeton, 1979), p. 193–238.

[17] J. Hadamard, J. Math. Pur. Appl. **4**, 27 (1898).

[18] G. A. Hedlund, Bull. Am. Math. Soc. **45**, 241 (1939).

[19] E. Hopf, Proc. Nat. Acad. Sci. **18**, 263 (1932).

[20] D. V. Anosov, Proc. Steklov Math. Inst. **90**, 1 (1967). See, for a review, Ya. G. Sinai, *Dynamical Systems II,* Encyclopaedia of Mathematical Sciences Vol. 2 (Springer-Verlag, Berlin, 1989).

[21] There are very interesting works on (i) the anisotropic Kepler problem, by M. C. Gutzwiller, J. Math. Phys. **18**, 806 (1977), and earlier works by the same author quoted therein; (ii) the periodic Coulomb potential, by A. Knauf, Commun. Math. Phys. **110**, 89 (1987). An application of Krylov's approach to the gravitational N-body problem can be found in V. G. Gurzadyan and G. K. Savvidy, Astron. Astrophys. **160**, 203 (1986).

[22] More details on the subject of this paragraph can be found in R. Abraham and J. E. Marsden, *Foundations of Mechanics* (Addison-Wesley, Redwood City, CA, 1987).

[23] G. S. Asanov, *Finsler Geometry, Relativity and Gauge Theories* (Reidel, Dordrecht, 1985).

[24] H. Rund, *The Differential Geometry of Finsler Spaces* (Springer-Verlag, Berlin, 1959).

[25] L. P. Eisenhart, Ann. Math. **30**, 591 (1939).

[26] A. Lichnerowicz, *Théories Relativistes de la Gravitation et de l'Electromagnétisme* (Masson, Paris, 1955).

[27] D. V. Alekseevskij, A. M. Vinogradov, and V. V. Lychagin, *Basic Ideas and Concepts of Differential Geometry,* Encyclopaedia of Mathematical Sciences Vol. 28 (Springer-Verlag, Berlin, 1991).

[28] The problem of geodesic spread was first discussed for Riemannian spaces by T. Levi-Civita, Math. Ann. **97**, 291 (1926), though the corresponding equation is commonly attributed to Jacobi, who studied the problem of second variation in variational calculus. Details on this equation can be found in M. Postnikov, *The Variational Theory of Geodesics* (Dover, New York, 1983).

[29] H. Fürstenberg and H. Kesten, Ann. Math. Stat. **31**, 457 (1960).

[30] G. Benettin, L. Galgani, and J. M. Strelcyn, Phys. Rev. A **14**, 2338 (1976).

[31] S. I. Goldberg, *Curvature and Homology* (Academic, New York, 1962).

[32] At heuristic level, this mechanism has been used in R. Lima and M. Pettini, Phys. Rev. A **41**, 726 (1990).

[33] E. A. Grebenikov and Yu. A. Ryabov, *Constructive Methods in the Analysis of Nonlinear Systems* (Mir, Moscow, 1983).

[34] R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani, J. Phys. A **18**, 2157 (1985).

[35] The classical Monte Carlo method with the so-called "importance sampling" was given in N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953). An adaptation of this algorithm to the microcanonical measure can be found in M. Creutz, Phys. Rev. Lett. **50**, 1411 (1983).

[36] R. Livi, M. Pettini, S. Ruffo, and A. Vulpiani, J. Stat. Phys. **48**, 539 (1987).

[37] A. M. Vinogradov, in *Mechanics, Analysis and Geometry: 200 Years after Lagrange,* edited by M. Francaviglia (Elsevier, New York, 1991).

[38] K. Yano and S. Ishiara, *Tangent and Cotangent Bundles* (Dekker, New York, 1973).

[39] S. Sasaki, Tohoku Math. J. **14**, 146 (1962).

[40] D. F. Escande, in *Large Scale Structures in Nonlinear Physics,* edited by J. D. Fournier and P. L. Sulem, Lecture Notes in Physics Vol. 392 (Springer-Verlag, Berlin, 1991).

[41] V. I. Arnold, C. R. Acad. Sci. Ser. A **261**, 17 (1965); M. Hénon, *ibid.* **262**, 312 (1966); H. Aref, J. Fluid Mech. **143**, 1 (1984).

[42] M. Pettini, A. Vulpiani, J. H. Misguich, M. De Leener, J. Orban, and R. Balescu, Phys. Rev. A **38**, 344 (1988).