# Relaxation and stochasticity in a truncated Toda lattice

Henry E. Kandrup* and M. Elaine Mahon†
*Department of Astronomy and Institute for Fundamental Theory, University of Florida, Gainesville, Florida 32611*
(Received 22 November 1993; revised manuscript received 14 February 1994)

This paper summarizes an investigation of the statistical properties of stochastic orbits in one fixed, time-independent potential, namely, the sixth order truncation of the Toda lattice potential, with the aim of identifying a meaningful notion of "relaxation," and then of correlating this relaxation with the overall stochasticity of the orbits. For variable energies $E$, localized ensembles of initial conditions were constructed and the subsequent evolution of these initial data then computed numerically. One discovers thereby that, at least above a critical energy $E_0$, most ensembles of stochastic orbits exhibit a rapid evolution, exponential in time, towards a time-independent invariant distribution, *not* microcanonical, the form of which is independent of the choice of the initial ensemble. Moreover, for fixed energy $E$, the decay rate $\Lambda$ associated with this exponential approach is independent of the size or location of the phase space region probed by the initial ensemble. This approach towards an invariant measure correlates directly with the sensitive dependence on initial conditions exhibited by the stochastic orbits: A small initial perturbation of the ensemble grows exponentially in time at a rate $\lambda$ which depends only on the energy $E$ and, within statistical uncertainties, the ratio $\Lambda(E)/\lambda(E)$ is independent of $E$. The fact that an initial ensemble of orbits evolves towards an invariant measure suggests that Lyapunov exponents computed for individual orbits should also have a physical meaning in terms of the shorter time evolution of ensembles of orbits. This intuition is corroborated by calculations that show that, in a well-defined sense, Lyapunov exponents $\chi(E)$ characterize the "average" instability of ensembles of orbits that sample the invariant measure.

PACS number(s): 05.90.+m, 51.10.+y, 98.10.+z

## I. INTRODUCTION

The research summarized herein has three principal aims, namely, (1) to formulate a simple statistical characterization of the short time evolution of a Hamiltonian system which evidences a considerable degree of stochasticity but is *not* completely ergodic; (2) to identify and quantify a notion of "relaxation towards a statistical (quasi) equilibrium" which is both physically useful and mathematically meaningful; and (3) to establish a direct correlation between this "relaxation" and the stochastic properties of the orbits.

The authors are especially interested in applications to astronomical systems, such as galaxies, where one is concerned with the orbits of individual stars, idealized as point masses, in the average gravitational potential associated with the overall distribution of mass. Such applications are interesting methodologically because of the fact that the characteristic time scales of interest can be extremely long. Thus, for example, the typical orbital period, $t_{cr}$, for stars like the Sun in a galaxy like the Milky Way is of order $2\times 10^8$ yr [1].

This implies that, if one focuses on a single stochastic orbit and computes a Lyapunov exponent, $\chi$, in the standard way, as the $t\rightarrow\infty$ limit of Lyapunov characteristic numbers $\chi(t)$ [2], she will typically be forced as a practical matter to integrate over periods of time that are orders of magnitude longer than $t_H$, the age of the Universe [3]. From the viewpoint of the nonlinear dynamicist, the evolution of orbits over time scales $\leq t_H$ involves short-term, transient behavior.

If the Hamiltonian system of interest is completely integrable, it is of course possible to provide a complete characterization of orbits in the potential, since the evolution is restricted simply to lower-dimensional hypersurfaces [4]. If the system is nonintegrable, this is no longer exactly so. Nevertheless, if most of the orbits are regular, as will be true for "nearly integrable" systems, there still exist various techniques which can be used to analyze and characterize the orbits, such as approximate integrals of the motion, classifications of the orbital types, and various spectral techniques [5].

However, if a substantial fraction of the orbits are stochastic, with positive Lyapunov exponents, these techniques are no longer applicable; and, moreover, there exists a sensitive dependence on initial conditions which renders a strictly deterministic analysis less obviously meaningful: As a practical matter, this exponential sensitivity implies that an experimentalist or observer cannot determine initial conditions with such precision as to justify a detailed deterministic evolution over even moderately long time scales; and, even if she could do so, it would not be possible, because of round-off error, etc., to generate numerically an accurate pointwise representation of the actual trajectory [6] (even though one might hope to justify numerical calculations in terms of some sort of "shadowing" argument [7]).

*Also at the Department of Physics, University of Florida, Gainesville, FL 32611. Electronic address: kandrup@astro.ufl.edu
†Electronic address: mahon@astro.ufl.edu

 ©1994 The American Physical Society

Suppose, for specificity, that one is interested in motion in a potential so chosen that the orbits are restricted to a compact phase space region, e.g., because $V$ is bounded from below and diverges as $|\mathbf{r}| \to \infty$. In this case, one might expect physically that, if the orbits are stochastic, so that two nearby trajectories typically diverge exponentially, they will tend quickly to "run all over" the accessible phase space on a time scale $\tau$ related somehow to the time scale on which nearby trajectories diverge. This "accessible phase space" might be the entire constant energy surface, or it might instead be a smaller region in which the orbits are confined by some additional constraints. These constraints could be either exact constraints, associated, for example, with invariant tori, or, alternatively, approximate constraints associated with Cantori through which orbits can "leak" only on time scales much longer than those of physical interest.

If one specifies some localized ensemble of initial conditions, and evolves these initial data into the future, there is no reason *a priori* to expect that, because of the coexistence of regular and stochastic regions, the ensemble will eventually evolve towards a uniform population of the constant energy surface, i.e., towards a microcanonical distribution. However, one might nevertheless expect that, at late times, different initial ensembles of stochastic orbits will, if they all have access to the same phase space region, tend, at least statistically, to sample the accessible phase space in a fashion that is independent of initial conditions. In other words, one might anticipate an evolution towards a time-independent invariant measure.

This possibility impacts directly on a problem of substantial interest for the gravitational $N$-body problem [8,9]. On the one hand, numerical simulations have shown that, when viewed in the many-particle phase space, the $N$-body problem is exponentially unstable towards small changes in initial conditions on a time scale comparable to a typical crossing time, $t_{\rm cr}$ [10]. On the other hand, simulations have also shown that, given generic initial data, an $N$-body system will typically evolve towards a statistical quasiequilibrium, in which bulk properties evidence only small variability, on the same time scale $t_{\rm cr}$ [11]. A critical question which one would like to answer is, can one demonstrate a precise sense in which this "relaxation" towards a statistical quasiequilibrium is directly related to the stochasticity of the orbits; i.e., to the exponential sensitivity?

In the preceding discussion, the notion of an invariant measure can be understood in two different contexts. More fundamentally, one can visualize a completely rigorous evolution towards an invariant measure which obtains exactly in the $t \to \infty$ limit. However, one can also visualize a heuristic notion of an "approximate" invariant measure, which may prove more useful for understanding evolution on shorter time scales. Suppose, for example, that one is interested in physical processes that are relevant on time scales $\sim t_H$. It would then be significant if, on a time scale $\ll t_H$, generic initial data were to exhibit an evolution towards a near-invariant measure which, albeit not strictly time independent, only evidences time-dependent changes on a time scale $\gg t_H$.

There is substantial numerical evidence that precisely this situation obtains for the gravitational $N$-body problem. Specifically, although numerical simulations exhibit an evolution towards a time-independent quasiequilibrium on a time scale $\sim t_{\rm cr} \ll t_H$, they also indicate that this quasiequilibrium is not a true equilibrium, and that, if one waits long enough ($t \gg t_H$), one will eventually see further evolutionary effects.

The aim of this paper is to address the problem of providing a short time statistical characterization of stochastic orbits and the problem of ascertaining whether, in the aforementioned sense, "relaxation" and stochasticity can be related. This will be done by answering several concrete questions about stochastic orbits in one particular model potential, namely, the sixth order truncation of the Toda lattice [12]. The focus here is not on asymptotic properties appropriate in a $t \to \infty$ limit. Rather, attention will be restricted primarily to the properties of orbits over times $t < 100$, this particular time interval corresponding in "physical" units to an astronomically relevant time scale $t_H$.

(1) For different localized ensembles of initial conditions, each corresponding to stochastic orbits, is there an approach towards an invariant measure, the form of which is independent of the specific choice of initial ensemble? As described below, the answer to this is yes. Moreover, the characteristic time scale $\tau$ associated with the approach towards the invariant measure is short. In "physical" units, typically $\tau \ll t_H$.

(2) For fixed energy $E$, is the time scale $\tau$ associated with this approach independent of the choice of initial conditions? The answer to this is a qualified yes. At relatively high energies, most initial ensembles evolve exponentially towards the invariant measure on the same time scale $\tau$, but there *are* some exceptional ensembles in which the approach is somewhat slower. Moreover, as the energy decreases, and the overall proportion of stochastic orbits decreases [13], the relative abundance of these "exceptions" increases.

(3) Is the approach towards such an invariant measure exponential in time $t$ or is it instead given as a power law in $t$? The answer here is that, for a broad range of energies $E$, the approach is exponential and, moreover, that the convergence rate $\Lambda(E)$ is well fit as a linear function of $E$, i.e., $\Lambda(E) = A + BE$, for positive constants $A$ and $B$. Especially in view of the aforementioned "exceptions," it is not clear whether there is a sense in which this general scaling persists for stochastic orbits at all energies, or whether, instead, it breaks down below some critical energy $E_0$.

(4) To what extent does this approach towards an invariant measure reflect the sensitive dependence on initial conditions? In particular, is there a simple correlation between (a) the time scale associated with the approach towards an invariant measure and (b) the time scale associated with the sensitive dependence on initial conditions, as probed, e.g., by the rate at which a small initial perturbation of the ensemble grows? Here again the answer is yes.

Suppose that each initial condition in an ensemble of initial conditions in a stochastic region is subjected to

some small perturbation in position and/or momentum, and the total configuration and momentum space perturbations for the ensemble,

$$\Delta r = \sum_{i=1}^{N} (\delta x_i^2 + \delta y_i^2)^{1/2} , \tag{1.1}$$

$$\Delta p = \sum_{i=1}^{N} (\delta p_{x,i}^2 + \delta p_{y,i}^2)^{1/2} , \tag{1.2}$$

are tracked as functions of time. One then concludes (a) that, until they become macroscopic, $\Delta r$ and $\Delta p$ grow exponentially, at a rate that is, for fixed energy $E$, independent of the initial ensemble and the choice of initial perturbation; (b) that the growth rate $\lambda$ is also well fit by a linear relation $\lambda(E)=a+bE$; and (c) that, within statistical uncertainties, the ratio $\Lambda/\lambda$ is approximately constant, independent of energy. There is a direct one-to-one correlation between the rate at which nearby ensembles diverge and the rate at which generic ensembles evolve towards an invariant measure. Similar correlations also exist between $\Lambda$ and other measures of the sensitive dependence on initial conditions, such as the Lyapunov exponent.

It is very convenient mathematically to characterize the overall exponential instability of stochastic orbits by computing a Lyapunov exponent, $\chi$. However, as observed already, the usual construction via a $t \to \infty$ limit is arguably unphysical for problems of astronomical interest in the sense that it entails following orbits for times that are much longer than $t_H$, the age of the Universe. It is more natural physically to consider a large number of orbits but only follow their evolution for time scales $\leq t_H$.

Fortunately, the fact that localized ensembles evolve exponentially towards an invariant measure on a relatively short time scale suggests a natural algorithm for estimating a Lyapunov exponent without studying individual orbits for extremely long times. The critical point is that, because the invariant measure is (at least approximately) time independent, the statistical properties of a single stochastic orbit, as evaluated over extremely long time scales, should coincide with the statistical properties of an ensemble of orbits, analyzed over shorter time scales, *provided that the initial ensemble is selected with a phase space weighting determined by the invariant measure.* The equivalence of temporal and phase space averages follows from Birkhoff's ergodic theorem [14] for a true invariant measure.

To compute a shorter time estimate of the Lyapunov exponent, all that one need do is generate initial conditions by sampling the invariant measure, compute shorter time Lyapunov characteristic numbers $\chi(t)$, and then construct an appropriate average. What this implies is that the Lyapunov exponent has a physical meaning even on relatively short time scales, by providing a precise characterization of the average instability of "random" orbits that sample the invariant measure. Finite time Lyapunov characteristic numbers have already been considered in other, related contexts by a number of different workers [15], and it has been shown that they can provide useful insights into various problems of physical interest [16], including fluid mechanics and plasma physics.

Section II of this paper provides a description of the experiments that were performed, and the fashion in which the data were analyzed. Section III then summarizes the results derived from the simulations involving the approach towards an invariant measure. Section IV summarizes the results of experiments in which small perturbations were introduced into the initial ensembles, demonstrating in particular how the total growth rates $\lambda(E)$ correlate with the decay rates $\Lambda(E)$ associated with the exponential approach towards an invariant measure. Section V concludes by showing that, in a precise sense, the standard Lyapunov exponents do indeed provide information about the "average" rate at which nearby trajectories diverge on relatively short, and hence physically relevant, time scales.

## II. A DESCRIPTION OF THE SIMULATIONS

The exact Toda lattice Hamiltonian, which corresponds to an integrable system with two degrees of freedom [12], is of the form

$$H = \tfrac{1}{2}(p_x^2+p_y^2) + \tfrac{1}{24}\{\exp[2\sqrt{3}x+2y] + \exp[-2\sqrt{3}x+2y] + \exp(-4y)\} - \tfrac{1}{8} , \tag{2.1}$$

where $\{x,p_x\}$ and $\{y,p_y\}$ represent conjugate pairs. The sixth order approximation, obtained by realizing the potential $V(x,y)$ as a Taylor series in $x$ and $y$ and then truncating at sixth order, involves a Hamiltonian

$$H = \tfrac{1}{2}(p_x^2+p_y^2) + V(x,y) , \tag{2.2}$$

where

$$V(x,y) = \tfrac{1}{2}(x^2+y^2) + x^2y - \tfrac{1}{3}y^3 + \tfrac{1}{2}x^4 + x^2y^2 + \tfrac{1}{2}y^4 + x^4y + \tfrac{2}{3}x^2y^3 - \tfrac{1}{3}y^5 + \tfrac{1}{5}x^6 + x^4y^2 + \tfrac{1}{3}x^2y^4 + \tfrac{11}{45}y^6 . \tag{2.3}$$

Unlike the exact Toda Hamiltonian, this $H$ defines a system which is *not* integrable and which, for energies above $E \approx 0.80$, admits stochastic orbits with positive Lyapunov exponent [13]. For energies between $E \approx 0.80$ and 24, the relative proportion of stochastic orbits increases rapidly to approximately 80%, but, for larger energies, the fraction of stochastic orbits becomes a much more slowly increasing function of $E$ (cf. Fig. 12 in Ref. [13]).

For different values of energy between $E=5$ and 75, surfaces of section were generated, plotting $y$ and $p_y$ at successive points where randomly chosen orbits pass through the value $x=0$. This is a useful choice of section because the potential (2.3) is symmetric under a reflection $x \to -x$, so that each orbit must periodically intersect the $x=0$ hyperplane. These surfaces of section were then used to identify initial data corresponding to stochastic orbits. Specifically, ensembles of stochastic orbits were generated by (a) uniformly sampling a rectangle in the $\{y,p_y\}$ plane which corresponds to stochastic orbits with $x=0$, and then (b) computing for each choice of $y$ and $p_y$

a value for $p_x$, assumed non-negative, by demanding that the orbit have the specified energy $E$, i.e., selecting phase space coordinates $\{0,y,p_x,p_y\}$ with

$$p_x \equiv +\{2[E-V(0,y)]-p_y^2\}^{1/2} . \tag{2.4}$$

Most of the orbit integrations involved ensembles comprised of 400 orbits. However, several integrations involved 1600 orbits and, in one case, 3200 orbits were considered. In absolute units, the total time over which the orbits were integrated was $t=100$. The orbits were integrated using a fourth order Runge-Kutta scheme with fixed time steps. The integrations at low energies involved a time step $h=10^{-3}$ but, for higher energies, $10^{-4}$ was used. The positions, momenta, and energy of each orbit were recorded at $t=0.25$ intervals.

As a test of the reliability of the simulations, several different ensembles were evolved using three different choices of time step, namely, $h=10^{-2}$, $10^{-3}$, and $10^{-4}$. It was found that, for relatively short time scales, the results of the three different integrations were essentially identical, but, for somewhat longer time scales, significant differences began to arise. However, except for the highest energies that were considered, namely, $E=60$ and 75, integrations with $10^{-3}$ and $10^{-4}$ for times $t \leq 100$ yielded no statistically significant differences in the quantities computed in Secs. III and IV. This was interpreted as a justification for using $h=10^{-3}$ times steps for the lower energies. It should also be stressed that, in all cases, the initial time interval used in computing the time scale $\tau$ associated with an approach towards the invariant measure was sufficiently short, $t \leq 40$, that the data could be trusted in a pointwise sense, orbit by orbit.

The data from the orbits at $t=0.25$ intervals were binned in the phase space variables, $x$, $y$, $p_x$, and $p_y$ to obtain coarse-grained approximations $F(x,y,p_x,p_y,t)$ of the time-dependent distribution function. Such an $F$ is, however, unwieldy. Because $F$ depends on four variables at each time $t$, it is difficult to visualize; and, even more importantly, because one is only considering $\leq 3200$ orbits, a gridding into $n^4$ cells with $n$ not extremely small would lead to statistically insignificant results. For this reason, attention focused primarily on "reduced" distribution functions $f(t)$ constructed by summing over two of the four phase space variables. There are of course six possible choices of reduced $f$. Attention focused primarily on three of these, namely, $f(x,y,t)$, $f(p_x,p_y,t)$, and $f(y,p_y,t)$.

The coarse-grained reduced distributions were all constructed by binning into a rectangular lattice of equal area cells, of total size $n \times n$, with $n=10$, 20, and 40. The region to be gridded varied as a function of $E$, as it was altered to ensure (1) that all points along each orbit fitted into the region that was gridded but (2) that only a minimum number of cells were forbidden energetically (because of the triangular symmetry of the potential, a rectangular gridding necessarily yields some inaccessible cells).

Such coarse-grained distributions $f(t)$ are still difficult to analyze since, e.g., for a $10\times 10$ binning of 400 orbits, at each instant of time this would entail assigning 400 points into 100 different cells, of which $>80$ are accessible. For this reason, an additional temporal coarse graining was effected by averaging $f$ over successive $t=0.25$ times steps. More precisely, a coarse-grained $\bar{f}(t)$ was defined by the prescription

$$\bar{f}(t)=\sum_{t}^{t+m\delta t} f(t) , \tag{2.5}$$

with $\delta t=0.25$ and $m=15$ or 25. It was discovered that the two different choices of $m$ lead to distributions that are statistically identical.

As discussed more carefully in Sec. III, it was observed that $\bar{f}(t)$ evolves exponentially on a relatively short time scale towards an invariant measure $f_{\rm inv}$, the form of which is, for fixed energy $E$, essentially independent of the choice of initial ensemble. A specific coarse-grained realization of this invariant $f_{\rm inv}$ was generated in the following way: For each choice of initial ensemble, compute the coarse-grained reduced $f(t)$ at $t=0.25$ intervals. Next, for the given ensemble, average $f(t)$ over the last 200 recorded values of $t$, i.e., from $t=50.25$ to 100.00. Finally, average the results obtained thereby for the different ensembles that were evolved. Via this prescription, eight ensembles of $N=400$ orbits yield an $f_{\rm inv}$ generated from $6.4\times 10^5$ points. This provides reasonable statistics, even for a $40\times 40$ gridding, which involves $1.6\times 10^3$ cells.

The form of the reduced invariant distribution was examined visually in two different ways, namely, (1) by constructing density contour plots from the gridded data and (2) by constructing gray-scale maps. One obvious question of interest is the extent to which the full invariant measure $F_{\rm inv}(x,y,p_x,p_y)$ corresponds to a microcanonical distribution $F_{\rm micro}$, i.e., to a uniform sampling of the constant energy surface. This is in fact straightforward to ascertain. Indeed, one verifies quite generally that, if the invariant measure is of the form

$$F_{\rm micro}=A\delta_D(H-E) , \tag{2.6}$$

where $A$ is a constant, $\delta_D$ denotes a Dirac distribution, and the Hamiltonian $H$ is of the form (2.2) for any potential $V(x,y)$, the reduced spatial distribution

$$f_{\rm micro}(x,y)=\begin{cases} 2\pi A & \text{if } x,y \text{ so that } V(x,y)\leq E \\ 0 & \text{if } x,y \text{ so that } V(x,y)>E \end{cases} . \tag{2.7}$$

(If, instead, one had a three-degree of freedom system, with a kinetic energy $K=p^2/2$, the constant term in Eq. (2.7) would be replaced by $f_{\rm micro}(\mathbf{r})=4\pi A\,[2(E-V)]^{1/2}$.) One can infer that $F_{\rm inv}$ deviates from a microcanonical distribution if $f_{\rm inv}(x,y)$ shows statistically significant differences from the piecewise constant distribution (2.7).

To quantify the sense in which $\bar{f}(t)$ converges towards $f_{\rm inv}$, one needs to identify a notion of "distance" between two coarse-grained reduced distribution functions, $f_1$ and $f_2$. This was done by constructing a discretized analog of an $L^1$ norm, which entails a pointwise comparison of the population of each cell of the coarse-grained distribution functions. Consider for specificity the reduced distribu-

tions $\bar{f}(x,y,t)$ and $f_{\rm inv}(x,y)$, each of which is assumed to have been normalized identically, so that $\sum_x \sum_y \bar{f}(x,y,t) = \sum_x \sum_y f_{\rm inv}$. The distance $Df(x,y,t)$ is then defined by the prescription

$$Df(x,y,t) = \frac{\sum_x \sum_y |\bar{f}(x,y,t) - f_{\rm inv}(x,y,t)|}{\sum_x \sum_y f_{\rm inv}(x,y,t)} . \quad (2.8)$$

Note in particular that $Df(x,y,t)=0$, i.e., $\ln Df \to -\infty$, implies complete agreement between the true and invariant distributions.

This definition provides a strong notion of convergences in that it involves a pointwise comparison of the two coarse-grained distributions. Mathematically, an $L^2$ norm [constructed in terms of $|\bar{f}(x,y,t) - f_{\rm inv}(x,y,t)|^2$] might seem more natural. However, one can argue that the prescription adopted here is physically better motivated because it is $f$, rather than $f^2$, which is the basic object that defines probabilities. The "pointwise" character of the comparison is of course diluted somewhat in that one is considering sums of coarse-grained distributions, rather than integrals of smooth functions. However, as will be discussed in Sec. III, it would appear that the rate at which $Df$ converges towards zero is in fact independent of the scale of the coarse graining.

The average rate at which nearby stochastic trajectories diverge was computed in three different ways. The first involved subjecting each individual orbit in the ensemble to some small change in position and momentum, and then tracking the evolution of the total configuration and momentum space perturbations, $\Delta r$ and $\Delta p$, until they become macroscopic. Explicit calculations show that perturbations of individual orbits typically grow exponentially at a rate that is largely independent of the detailed form or amplitude of the initial perturbation. It is therefore natural, and convenient, to compute $\Delta r$ and $\Delta p$ by perturbing each initial condition in the ensemble in an identical fashion. This was done by perturbing each initial condition with specified $y$ and $p_y$ to new values $y+0.0001$ and $p_y+0.001$, with $x$ still assumed to vanish identically and $p_x$ modified to ensure that the energy $E$ is unchanged.

This statistic provides a simple characterization of the overall rate at which small perturbations grow, but it is limited in that it only probes the aggregate properties of a large number of orbits over the relatively short period of time before the perturbation becomes large. For this reason, the average rate at which nearby stochastic trajectories diverge was also computed in two other ways, both of which involve the standard prescription due to Bennetin, Galgani, and Strelcyn [2]. Over finite times $t$ one can define a Lyapunov characteristic number

$$\chi(t) = \lim_{\delta r(0)\to 0} \frac{1}{t} \frac{\delta r(t)}{\delta r(0)} , \quad (2.9)$$

where $\delta r(0)$ and $\delta r(t)$ denote the configuration space deviations of two nearby orbits at time 0 and $t$, and, in the limit $t \to \infty$, one obtains thereby the Lyapunov exponent

$$\chi = \lim_{t\to\infty} \chi(t) . \quad (2.10)$$

It follows from Oseldec's theorem [17] that different stochastic orbits all restricted to the same phase space region will be characterized by the same Lyapunov exponent. However, the time scale on which any individual orbit manages to probe the entire phase space region will in general be very long; and, for this reason, one might also expect that, as a practical matter, the time required to compute a reasonable estimate of the late times Lyapunov exponent $\chi$ from a single orbit will also be very long. These expectations, realized by other workers for other potentials, were also confirmed by the calculations described in Sec. IV. Even an evaluation of $\chi(t)$ for a time $t=10^4$ leads to overall uncertainties of amplitude $\sim 10\%$. The Lyapunov exponent $\chi$ most certainly does *not* provide a reasonable characterization of the instability of individual orbits on times scales $t \leq 100$.

The calculations of Lyapunov exponents from the long time evolution of individual orbits were therefore augmented by additional calculations which tracked $\chi(t)$ for multiple orbits for shorter total times $t_{\rm fin} \leq 100$. For each energy $E$, a fixed number $q$ of initial conditions were selected "at random," with a relative density weighted by the invariant distribution $F_{\rm inv}(x,y,p_x,p_y)$. These initial data were then evolved to derive characteristic numbers $\chi_i(t_{\rm fin})$ for each orbit $i$. Finally, the results from the individual orbits were averaged to derive a "mean Lyapunov number"

$$\bar{\chi} = \frac{1}{q} \sum_{i=1}^{q} \chi_i(t_{\rm fin}) \quad (2.11)$$

for the ensemble.

## III. THE APPROACH TOWARDS AN INVARIANT MEASURE

The orbit integrations described in Sec. II lead to several concrete conclusions regarding the evolution of an initially localized ensemble of stochastic orbits of fixed energy $E$.

The first unambiguous conclusion is that, at least for energies $E > 10$ or so, initially localized ensembles of stochastic orbits tend generically to evolve at late times towards a time-independent distribution. This was confirmed directly through the observation that all six coarse-grained reduced distributions $\bar{f}(t)$ evolve towards time-independent values $f_{\rm inv}$. Strictly speaking, this does not prove rigorously that the full $F(x,y,p_x,p_y,t)$ evolves towards an invariant $F_{\rm inv}$. However, what is true is that $F(x,y,p_x,p_y,t)$ must itself evolve towards an invariant distribution unless it incorporates finely tuned time-dependent correlations between the four basic variables.

The second unambiguous conclusion is that, at least for energies $E \geq 30$, the form of the invariant distributions $f_{\rm inv}$ is independent of the location of the initial ensemble. Different ensembles of orbits, originally situated in very different phase space regions, diverge in such a fashion as to sample the same measure at late times. That this result can be true is a reflection of the fact that, at sufficiently high energies, all the stochastic orbits have access to the same phase space region. At lower energies, this is no longer so. Specifically, one finds that, at least

for $E \leq 20$, a substantial fraction of the stochastic orbits divide into three distinct "families" which, at least on relatively short time scales $t \leq 100$ do not access the entire stochastic region.

This general behavior is illustrated by Figs. 1 and 2. Figure 1 plots the spatial coordinates $x$ and $y$ for one initially localized ensemble of 25 obits with $E=30$ at $t=0.25$ intervals for a total time $t=50$. The obvious point is that the orbits appear to fill a region in configuration space corresponding approximately to an equilateral triangle. In particular, the overall distribution manifests the discrete $2\pi/3$ rotational symmetry of the potential given by Eq. (2.3). Figure 2 plots $x$ and $y$ for a localized ensemble of 25 orbits with $E=10$, again at $t=0.25$ intervals and for a total time $t=50$. The obvious point now is that the region occupied by the orbits is no longer approximated by an equilateral triangle, and no longer manifests a $2\pi/3$ rotational symmetry. Most of the orbits are restricted instead to a smaller wedge in configuration space. For $E=30$, most/all ensembles of orbits yield plots analogous to Fig. 1. For $E=10$, some ensembles again yield plots with the same symmetries as Fig. 1. However, many yield instead plots like Fig. 2 or analogous plots derived by rotations of $\pm 2\pi/3$.

It is evident from Fig. 2 that some of the orbits "escape" from the localized wedge. Significantly, however, these escapers are not located near the edges of the localized ensemble, so that they cannot be eliminated by simple choosing an initial ensemble localized in a smaller phase space region. If one reduces the size of the region sampled by the initial ensemble, one does not completely eliminate the escapers. Rather, what appears to be true is that embedded in a phase space region in which "most" orbits are confined, at least on short time scales, are other orbits which are in fact able to escape. The fact that this conclusion appears to hold even for very small phase space regions suggests strongly that whether or not some orbit can escape from the localized wedge within a given

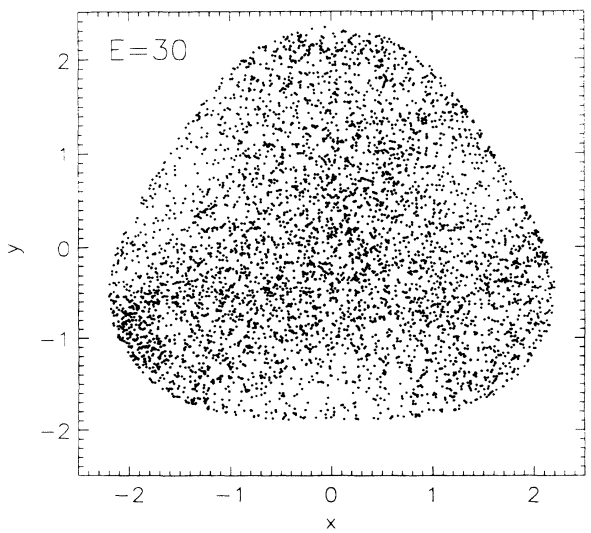


FIG. 1. The coordinates $(x,y)$ generated from 25 stochastic orbits with energy $E=30$, sampled at $t=0.25$ intervals for a total time $t=50$.

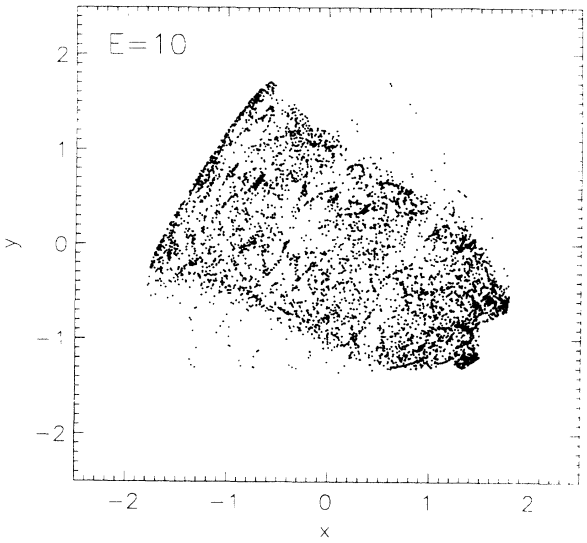


FIG. 2. The coordinates $(x,y)$ generated from 25 stochastic orbits with energy $E=10$, sampled at $t=0.25$ intervals for a total time $t=50$.

time interval $\Delta T$ evidences a fractal dependence on initial conditions (cf. [18]).

A third unambiguous conclusion is that the approach towards an invariant distribution, as probed through the "distance" $Df(t)$, is exponential. Analysis of the coarse-grained data demonstrates that a plot of $\ln(Df)$ as a function of $t$ is essentially linear. One *does* observe noticeable deviations from a pure exponential, deviations that become somewhat larger if one considers lower energies and/or ensembles localized initially in progressively smaller phase space regions. However, overall the approach towards an invariant distribution is well fit by an exponential relation

$$Df(t)=\text{const}\times\exp(-\Lambda t)\ . \tag{3.1}$$

A fourth unambiguous calculation is that the time scale $T=1/\Lambda$ associated with the approach towards an invariant measure is independent of the location of the initial ensemble of orbits. Different localized ensembles of initial conditions, situated in widely separated phase space regions, evolve towards the same invariant measure on the same time scale $T$. This behavior is illustrated in Fig. 3, which plots $\ln Df(x,y,t)$ for several different ensembles at $E=40$. For each individual ensemble, the data were constructed from coarse-grained reduced distribution functions $\bar{f}(x,y,t)$, obtained via a $10\times 10$ configuration space binning and a temporal average over $m=15$ time steps. The initial ensembles were each constructed via a uniform sampling of a phase space region with $\Delta y=0.2$ and $\Delta p_y=0.72$, assuming that $x=0$ and $p_x$ is given by Eq. (2.4). It is clear from Fig. 3 that the convergence towards an invariant measure is approximately exponential and that the time scale $T$ is approximately the same for each of the ensembles.

The data displayed in Fig. 3 all involve the particular reduced distribution $\bar{f}(x,y,t)$. This, however, is immaterial. A fifth unambiguous conclusion is that the time

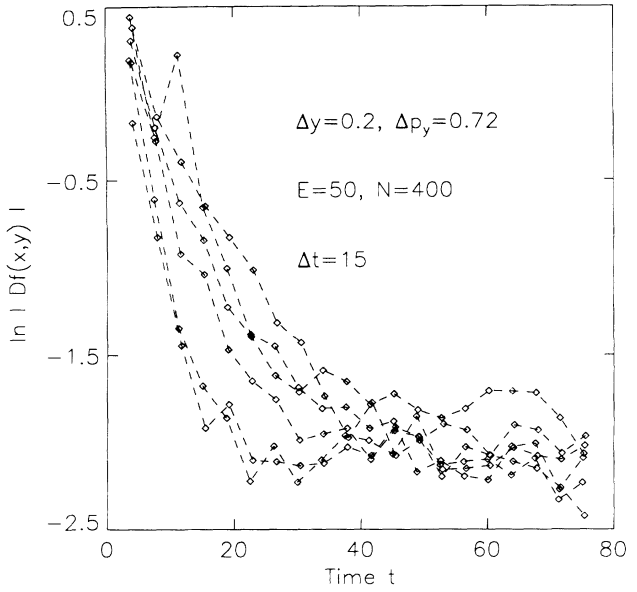


FIG. 3. $Df(x,y,t)$, the difference between $\bar{f}(x,y,t)$ and the invariant measure $f_{\text{inv}}(x,y)$, for several different ensembles of orbits with $E=40$. $Df$ was generated from a distribution $\bar{f}(x,y,t)$ which involved a $10\times10$ configuration space binning and a temporal averaging over 15 time steps.

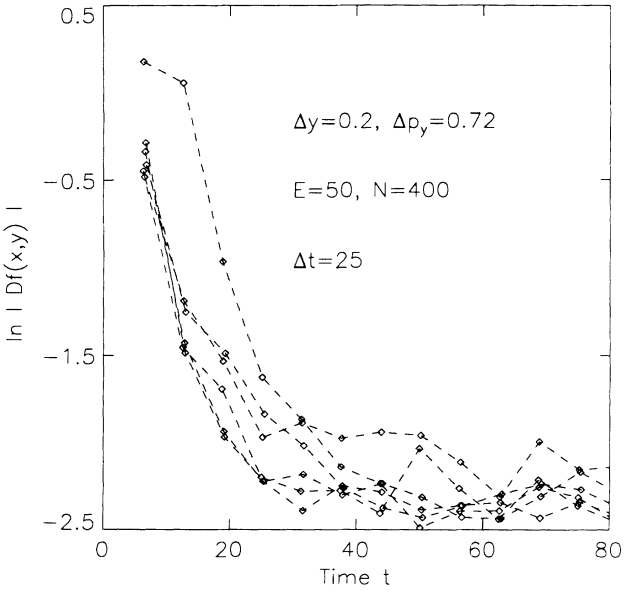


FIG. 4. The same as Fig. 3, except generated by a temporal averaging over 25 time steps.

scale $T$ associated with the approach towards an invariant $f_{\text{inv}}$ is independent of which pair of variables one chooses to consider, $\{x,y\}$, $\{x,p_x\}$, $\{x,p_y\}$, $\{y,p_x\}$, $\{y,p_y\}$, or $\{p_x,p_y\}$.

A sixth unambiguous conclusion is that, at least over a moderate range, the time scale $T$ associated with the approach towards an invariant distribution is independent of the level of the temporal and/or phase space coarse graining. Different temporal coarse grainings were effected by averaging over $m=15$ and 25 time steps, and it was observed that, within statistical uncertainties, these different coarse grainings did not change the basic conclusions. This is illustrated by Fig. 4, which is identical to Fig. 3 except that it was constructed from coarse-grained distributions $\bar{f}(x,y,t)$ that were averaged over $m=25$ time steps.

Similarly, coarse-grained distributions were constructed, allowing for $n\times n$ binnings, with $n=10$, 20, and 40, and it was discovered analogously that the level of this coarse graining does not alter the value of $T$. This is illustrated by Fig. 5 which, for one specific ensemble of initial conditions with $E=60$, computes $Df(x,y)$, $Df(p_x,p_y)$, and $Df(y,p_y)$ for the three different levels of phase space coarse graining. All the curves in this figure involve a distribution $\bar{f}(t)$ averaged over $m=15$ time steps. It is clear that, for early times, the slopes of the nine different plots in Fig. 5 are all essentially the same.

It is also evident that, at later times, the exponential decrease in $Df$ stops and the curves asymptote towards limiting values. This is a finite size effect. Neither the invariant distributions $f_{\text{inv}}$ nor the true coarse grained distributions $\bar{f}(t)$ associated with the initial ensemble have been determined exactly, since they have been constructed using only a finite number of orbits. Even if the numerical $\bar{f}(t)$ and $f_{\text{inv}}$ constitute two random samplings of the same analytic distribution, finite number statistics imply that $Df(t)$ must be nonzero. When these statistical errors become the dominant source of differences between $\bar{f}(t)$ and $f_{\text{inv}}$, the exponential decrease in $Df$ must stop.

To confirm explicitly that this "saturation" is purely a finite size effect, with no further physical implications, one can investigate how the time evolution of $Df(t)$

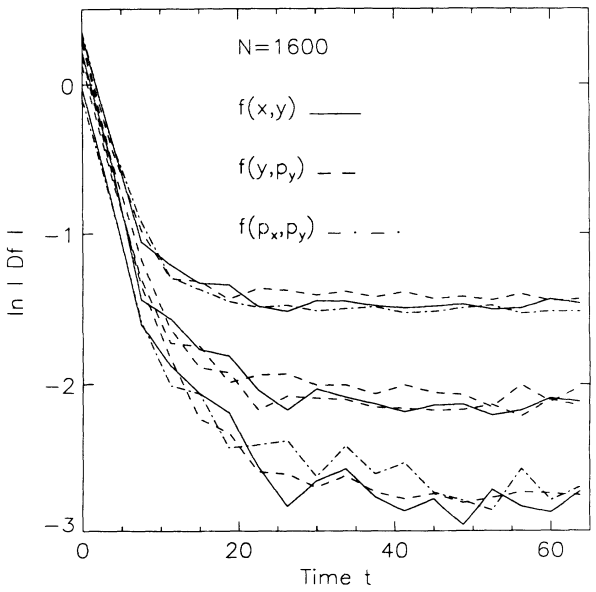


FIG. 5. $Df(x,y,t)$, $Df(y,p_y,t)$, and $Df(p_x,p_y)$ for one specific ensemble of initial conditions with $E=60$ with variable phase space coarse graining. The top three curves have a $40\times40$ coarse graining, the middle three have $20\times20$, and the bottom three have $10\times10$.

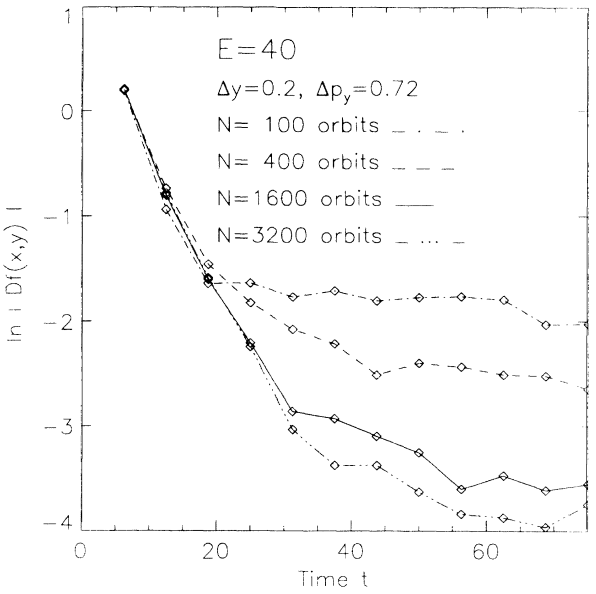


FIG. 6. $Df(x,y,t)$ computed for different ensembles of orbits with $E=40$, sampling the same phase space region but with variable numbers of orbits from $N=100$ to 3200.

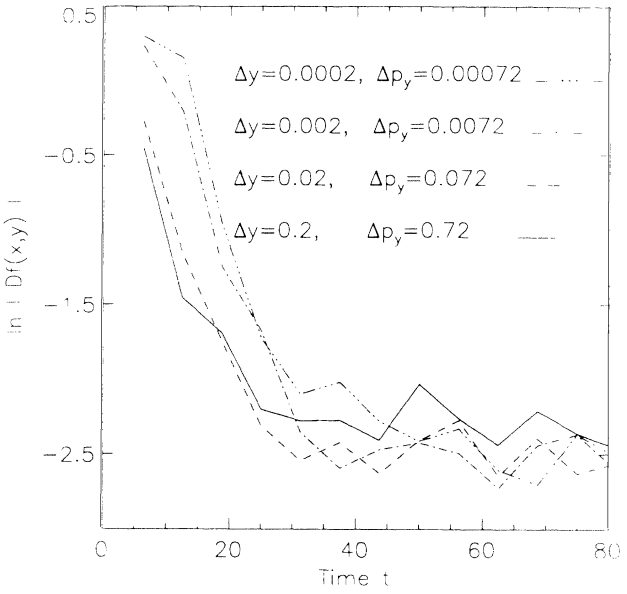


FIG. 7. $Df(x,y,t)$ computed for four different ensembles with $E=40$ centered about $(y,p_y)=(0,0)$. The different ensembles involve cells of initial conditions successively smaller in area by factors of 100.

varies as a function of total orbit number $N$. The results of one such investigation are exhibited in Fig. 6, which illustrates $\ln Df(t)$ computed for different ensembles of orbits with $E=40$, all sampling the same phase space region but allowing for variable $N$ from 100 to 3200. Note that, as expected, increasing $N$ lowers the value of $Df$ at which saturation occurs, so that the exponential decrease holds for longer times $t$.

A seventh unambiguous conclusion is that the time scale $T$ associated with the approach towards an invariant distribution is independent of the size of the phase space region sampled by the initial localized ensemble. This was examined in special detail for the case of orbits with $E=40$, allowing for $\{y,p_y\}$ phase space regions of variable size $0.2\times0.72$, $0.02\times0.072$, $0.002\times0.0072$, and $0.0002\times0.000\,72$, i.e., allowing for a systematic magnification in resolution by a factor of $10^6$.

Figure 7 exhibits the evolution of $Df(t)$ for four different localized ensembles, obtained by successive magnifications of a given phase space region with $E=40$ centered about $(y,p_y)=(0,0)$. Three facts follow from this and other similar sets of experiments. First of all, it is clear that, overall, the time required to converge towards an invariant measure increases as the size of the initial phase space region shrinks. Secondly, one discovers often (albeit not for the experiments summarized in Fig. 7) that the convergence towards the invariant measure becomes less uniform as the size of this region shrinks: one begins to see occasional deviations from a completely systematic decrease in $Df$. Thirdly, however, it is also clear that, despite these trends, the overall approach towards an invariant measure is still exponential and the time scale $T$ is essentially the same for the different magnifications. This is manifest in the fact that, although the curves in Fig. 7 have different offsets, at least initially their slopes are approximately the same.

An eighth conclusion is that, over a fairly wide range of energies $E$, the convergence time $T(E)$ and the convergence rate $\Lambda(E)$ are well fit by a simple analytic form. Specifically, one finds that, at least for energies in the range $10\leq E\leq 75$, the convergence rate is well fit by a linear growth law

$$\Lambda(E)=A+BE\ , \tag{3.2}$$

with $A=0.0330\pm0.0106$ and $B=0.002\,13\pm0.000\,48$. Because of the relatively large error bars, one cannot exclude the possibility of a more complicated functional dependence, but there is no evidence of any systematic deviations from the simple linear relation.

Another striking feature is that the late times distribution $F_{\rm inv}(x,y,p_x,p_y)$ is *not* microcanonical. This is, for example, illustrated by the form of the reduced distribution $f_{\rm inv}(x,y)$ which evidences statistically significant differences from the piecewise constant form implied by Eq. (2.7). At high energies, the distribution $f_{\rm inv}(x,y)$ exhibits four distinct local maxima, one at the origin, a second along the $y$ axis for a positive value of $x$, and two others obtained from the second by $\pm2\pi/3$ rotations. This, is, e.g., illustrated in Fig. 8, which presents a gray-scale plot of $f_{\rm inv}(x,y)$ for $E=30$. Note that Fig. 8 was generated from a large collection of 1600 initial conditions, including the 25 which were used in generating Fig. 1.

At lower energies, the situation becomes more complicated because of the existence of "families" of stochastic orbits which individually break the $2\pi/3$ rotational symmetry. However, even if one restricts attention to orbits which fill the entire triangular region, a qualitative change can be observed. Specifically, one discovers that

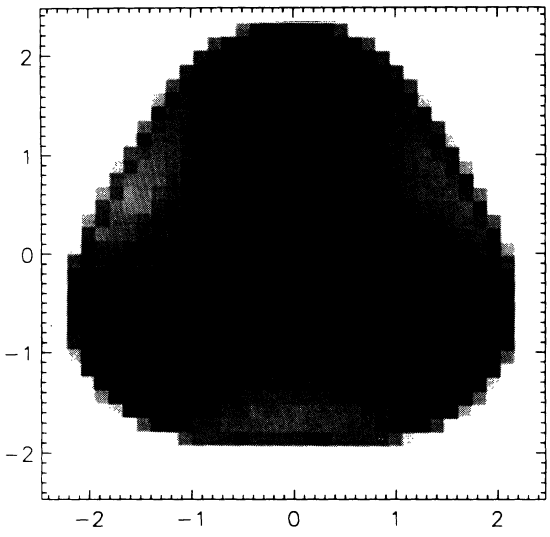


FIG. 8. A gray-scale plot of the projected invariant distribution $f_{\text{inv}}(x,y)$ for stochastic orbits with $E=30$.

the origin, which by symmetry must be a density extremal, eventually passes through a critical point and changes from a maximum to a local energy *minimum*. This is illustrated in Fig. 9, which presents a gray-scale plot of $f_{\text{inv}}(x,y)$ for $E=10$. This figure differs qualitatively from Fig. 2, also generated for orbits with $E=10$, because of the fact that the initial conditions correspond to orbits that are not "trapped" in a restricted portion of the accessible configuration space.

The aforementioned conclusions all refer to "generic" behavior. However, as observed already, exceptional ensembles were occasionally observed. Typically, these correspond to ensembles of stochastic orbits which initially begin to evolve exponentially towards the invariant measure, but then halt their approach, in the sense that $Df(t)$ stops decreasing exponentially towards zero before finite number effects have become important. This qualitative behavior is observed more frequently at lower energies, especially $E \leq 20$, and for initial ensembles that sample smaller phase space regions. Moreover, a qualitative inspection of the evolution of these ensembles would suggest that, typically, this behavior corresponds to a situation in which some fraction of the orbits have become trapped in a particular phase space region or in which the orbits tend systematically to avoid some given region.

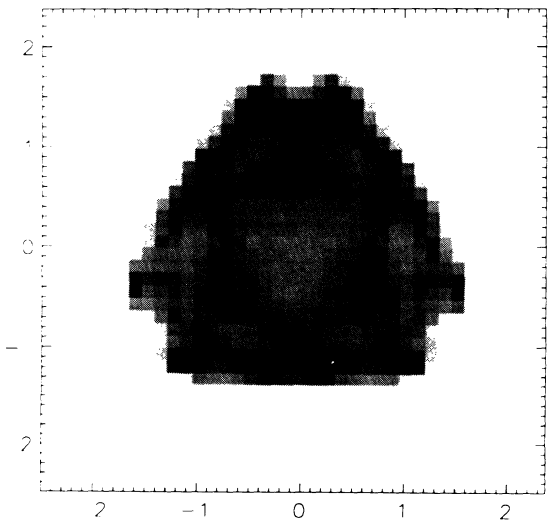


FIG. 9. A gray-scale plot of the projected invariant distribution $f_{\text{inv}}(x,y)$ for stochastic orbits with $E=10$.

It would seem natural to conjecture that this anomalous behavior is somehow associated with "sticky" islands: if some fraction of the orbits in the ensemble stray too close to an island, they may be temporarily detained in the neighborhood of the island and hence be prevented, at least for a time, from evolving towards an invariant measure. As $E$ decreases, the proportion of regular orbits, and hence of islands, increases and indeed, it is just below $E \approx 24$ that the relative abundance of stochastic orbits begins to decrease precipitously. Moreover, one may expect that smaller ensembles are more likely to be impacted significantly by tiny, seemingly randomly distributed islands in what appears superficially to be a sea of stochasticity. This interpretation, consistent with the interpretation of similar phenomena observed in other sorts of problems [18], would imply that most/all ensembles are effected at least somewhat by islands, but that, in many cases, the effects are simply too small to be observed.

In this connection, it should also be observed that, at energies $E < 10$ or so, the overall proportion of stochastic orbits and the size of typical simply connected stochastic regions are sufficiently small that it is difficult to construct large ensembles of stochastic orbits; and, for this reason, it is not clear numerically whether the simple behavior observed for ensembles of stochastic orbits at higher energies actually persists.

The computations summarized above show that localized ensembles of initial conditions corresponding to stochastic orbits exhibit a rapid exponential evolution towards a distribution which is rather nearly time independent. However, they do not, and cannot, prove that there are no further systematic changes on much longer time scales. One cannot, for example, exclude the possibility that, on very long time scales $t \gg 100$, the ensemble will evolve towards a distribution which is more nearly microcanonical.

The possibility of such a slow variability was in fact investigated for one set of initial conditions with $E=30$. Specifically, an initially localized ensemble of 1600 stochastic initial conditions was evolved into the future for a total time $t=1100$, and the output for the time interval [1050.25,1100] analyzed in the same fashion as the data for the interval [50.25,100] to extract new "invariant distributions" $f_{\text{inv}}(x,y)$, $f_{\text{inv}}(y,p_y)$, and $f_{\text{inv}}(p_x,p_y)$. This period of time is so long that one cannot trust the pointwise evolution of the orbits. However, one may hope that, because of some shadowing argument, the statistical properties of the orbits are reliable, and that this integration provides at least an indication of the qualitative form of the true evolution. In any event, a comparison of the early and late times "invariant distributions" suggests that the two different distributions may be slightly

different, but these differences, if real, are not of unambiguous statistical significance. One cannot exclude the possibility of nontrivial evolutionary effects on time scales $t \gg 100$. However, what seems clear is that any systematic evolutionary effects that do exist must proceed on an exceedingly long time scale.

## IV. THE SENSITIVE DEPENDENCE OF INITIAL CONDITIONS

Turn now to a consideration of the sensitive dependence on initial conditions that is exhibited by the stochastic orbits. As discussed in Sec. II, one especially simple way in which to quantify this sensitivity is by introducing small perturbations in the initial position and momentum of each member of an ensemble of stochastic orbits, and then computing the evolution of the total configuration and momentum space perturbations, $\Delta r$ and $\Delta p$, at subsequent times. The net result of such an investigation is that, as might have been expected, $\Delta r$ and $\Delta p$ typically grow exponentially, with a characteristic growth rate $\lambda$ that is independent of the specific form and amplitude of the initial perturbation, provided only that the perturbations of the individual orbits are all of comparable size.

This point was tested explicitly for the case of ensembles of orbits with $E=20$, where several different integrations were effected. Each of these experiments involved perturbing every initial condition in the same way, through given displacements $\delta x$, $\delta y$, and $\delta p_y$, i.e., $\{x,y,p_y\} \rightarrow \{x+\delta x, y+\delta y, p_y+\delta p_y\}$, with $p_x$ modified to guarantee that the energy is unchanged. Different experiments involved different choices of the initial displacements $\{\delta x, \delta y, \delta p_y\}$. The principal conclusions derived from these experiments were (1) that $\Delta r$ and $\Delta p$ typically grow exponentially until the initial perturbation becomes "macroscopic," at which point the overall growth saturates; (2) that the growth rates $\lambda_r$ and $\lambda_p$ associated, respectively, with $\Delta r$ and $\Delta p$ are, for each given experiment, equal to one another within statistical uncertainties; and (3) that the values of $\lambda_r = \lambda_v \equiv \lambda$ are independent of the specific choice of the initial perturbation, again to within statistical uncertainties.

The fact that the choice of initial perturbation is immaterial facilitates a simple algorithm which was used to quantify the sensitive dependence on initial conditions: Each ensemble analyzed in Sec. III was first perturbed via an initial displacement $y \rightarrow y+0.0001$ and $p_y \rightarrow p_y+0.001$, with $x$ still assumed to vanish and $p_x > 0$ fixed by the constraint of fixed energy. The perturbed ensembles were then evolved into the future, and the output of the perturbed ensembles compared with the output of the unperturbed ensemble to extract $\Delta r$ and $\Delta p$.

One concrete conclusion derived from these experiments is that, for fixed energy $E$, $\Delta r$ and $\Delta p$ do indeed grow exponentially with a growth rate $\lambda$ that is largely independent of the location of the initial ensemble. As for the case of the approach towards an invariant measure, some occasional exceptions were observed, these corresponding to a situation in which the perturbation initially grows somewhat more slowly than the average rate $\lambda$. However, these occasional anomalies only persist for relatively short periods of time, (say) $\Delta T \sim 2-3$, and, if one ignores them, one finds that the overall growth rate during the "normal" period is again given by the average $\lambda$, at least to within statistical uncertainties.

The second concrete conclusion is that the growth rate $\lambda$ varies smoothly as a function of $E$ in a fashion again consistent with a linear growth law. Specifically, the data are reasonably well fit by a linear relation

$$\lambda(E) = a + bE \ , \tag{4.1}$$

where $a = 0.2493 \pm 0.0404$ and $b = 0.0141 \pm 0.0009$. Careful examination of the data suggests that the $\lambda$ versus $E$ relation actually exhibits some curvature, especially at lower energies, but the deviations from a linear relation are small.

It is illuminating to compare the growth rates $\lambda$ with the decay rates $\Lambda$ associated with the evolution towards an invariant measure, which are also approximated by a linear relation, $\Lambda = A + BE$. The third concrete conclusion derived from such a comparison is that the ratio $R(E) \equiv \Lambda/\lambda$ is rather nearly constant, independent of energy $E$. Specifically, one finds a best fit value $R = 6.790 \pm 0.756$. The goodness of fit is illustrated by the upper curve in Fig. 10. The large error bars, especially prominent at lower energies, imply that one cannot exclude the possibility of relatively small systematic effects. However, what *is* clear is that $R$ does not change all that drastically.

This fact implies that, for stochastic orbits in this truncated Toda potential, there is a direct, one-to-one correspondence between the degree of stochasticity exhibited by an ensemble of orbits and the rate at which that ensemble evolves towards an invariant measure. Both the degree of stochasticity, as measured by the total growth rate $\lambda$, and the rate $\Lambda$ associated with the evolu-

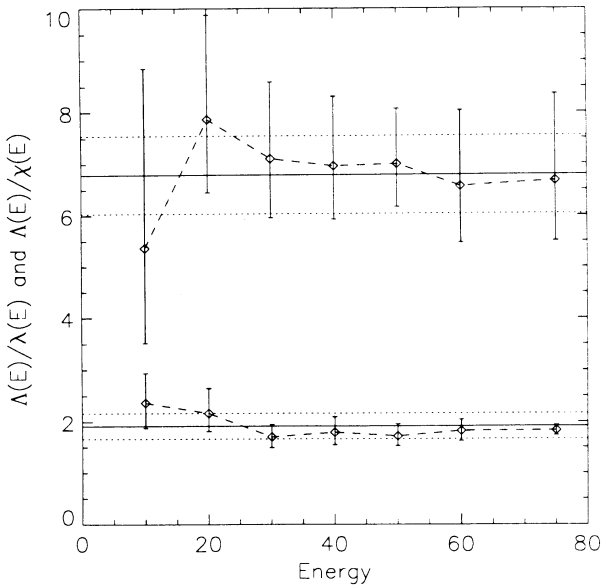


FIG. 10. The top curve exhibits the ratio $R = \Lambda(E)/\lambda(E)$ as a function of $E$. The lower curve exhibits the ratio $\mathcal{R} = \Lambda(E)/\chi(E)$ as a function of $E$.

tion towards an invariant measure are, for fixed energy $E$, essentially independent of the specific choice of initial ensemble. And, moreover, if $E$ is altered, $\lambda$ and $\Lambda$ change in such a fashion that the ratio $\Lambda/\lambda$ is rather nearly unchanged.

## V. SHORT TIMES ESTIMATES OF LYAPUNOV EXPONENTS

The stability of stochastic orbits in the truncated Toda potential was also examined in a more conventional way (cf. Ref. [2]) through a computation of Lyapunov exponents. This was done by introducing a small initial perturbation $\delta x = 10^{-10}$ into each orbit, continually solving the variational equations for the perturbation as the unperturbed orbit is evolved, and then renormalizing the evolved perturbation back to a total amplitude

$$(\delta x^2+\delta y^2+\delta p_x^2+\delta p_y^2)^{1/2}=10^{-10} \tag{5.1}$$

at intervals $\Delta t = 10$. Successive estimates for $\chi(t)$ were effected at the same $\Delta t$ intervals, with the calculations of an individual orbit proceeding for a total time $t = 10^4$. The value $\chi(t=10^4)$ was then interpreted as providing an estimate of the true Lyapunov exponent, which is only defined in a $t \to \infty$ limit. Several orbits were tracked for a period twice as long, namely, $t = 2\times 10^4$, and it was found that, overall, doubling the integration time had a relatively small effect on the final value of $\chi$. To test the reliability of the estimated Lyapunov exponent derived from a single orbit, computations of $\chi(t=10^4)$ were effected for a total of ten stochastic orbits for each value of $E$. An analysis of the means and the associated standard deviations implies that an integration for any individual orbit over a period $t = 10^4$ provides an estimate of $\chi$ accurate at approximately the 10% level.

For variable energies $E$, Fig. 11 exhibits as triangles the "average" Lyapunov exponents computed in this fashion, with the associated dispersions represented as error bars. It is evident that, unlike the instability time scale $\lambda(E)$ considered in Sec. IV, which can be reasonably well fit by a linear relation, $\lambda = a + bE$, the exponent $\chi(E)$ is *not* well fit as a simple linear function of $E$. Empirically, the data in Fig. 11 are better fit by a quadratic function $\chi(E)=\alpha+\beta E+\gamma E^2$, where $\alpha=-0.0148$, $\beta=0.0159$, and $\gamma=-8.82\times10^{-5}$. A naive least squares fit to a *linear* relation would yield instead $\chi(E)=\mathcal{A}+\mathcal{B}E$, with $\mathcal{A}=0.103$ and $\mathcal{B}=0.008\,53$, but it is evident that neither this nor any other linear relation can provide a reasonable fit to all the points. Moreover, even an empirical quadratic fit ignores a definite indication of a change in the functional form of $\chi(E)$ at energies below a critical value $E \approx 30$.

Despite these complications, however, there is still a direct connection between the Lyapunov exponent $\chi$ and the decay rate $\Lambda$ associated with the exponential approach towards an invariant measure. This is illustrated by the lower curve in Fig. 10, which exhibits the ratio $\mathcal{R}(E)=\Lambda/\chi$. There is an indication that, for small $E$, $\mathcal{R}$ is larger than for large $E$ but, overall, the data are consistent with a ratio that is approximately constant. The best fit value $\mathcal{R}=1.912\pm0.252$.

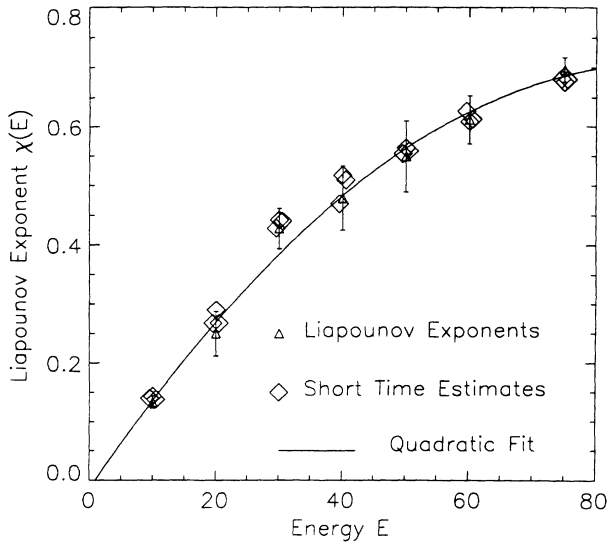


FIG. 11. Lyapunov exponents $\chi$ for orbits of varying energy $E$. The triangles represent mean values obtained by computing $\chi(t)$ for ten orbits for a total time $t = 10^4$. The error bars represent dispersions about these mean values. The diamonds represent mean values by computing $\chi(t)$ for 400 orbits for a total time $t = 100$. The solid curve is a least squares quadratic fit.

Estimates of the Lyapunov exponents were also generated by following for shorter times the evolution of orbits that sample the invariant measure. This was done by exploiting the data analyzed in Sec. III. Specifically, for each value of $E$, three of the 400 orbit ensembles were selected, and the final output at time $t = 100$ were chosen as initial data for a new ensemble, the presumption being that each of these sets of initial conditions should yield a fair sample of the true invariant measure. Estimates of $\chi(t)$ for each orbit in these new ensembles were then computed as above at time intervals $\Delta t = 10$, the only difference being that the integrations were halted after a much shorter total time, namely, $t = 100$.

The collection of 400 $\chi$'s at each time $t$ yields a *distribution of instability time scales,* for which a mean $\bar{\chi}(t)$ and a standard deviation $\sigma_\chi(t)$ were then computed. The net result of such a computation was the discovery that, in every case, the mean $\bar{\chi}(t=100)$ agreed with the conventional estimate of $\chi$ obtained from a long time integration of ten stochastic orbits. The goodness of fit is again illustrated in Fig. 11, where the mean values $\bar{\chi}(t=100)$ for the three different ensembles are represented by diamonds.

The basic conclusion is that, as one might have expected, the Lyapunov exponent $\chi$ can be interpreted physically as representing the "average" instability of an ensemble of orbits that samples the invariant measure. In other words, the fact that ensembles of stochastic orbits evolve towards an invariant measure, where temporal and phase space averages agree, implies that Lyapunov exponents, defined formally via $t \to \infty$ limit for a single stochastic orbit, also have a well-defined physical meaning for ensembles of orbits on much shorter time scales.

Although the mean $\bar{\chi}$'s agree closely with the long time estimates of $\chi$, the overall distribution of short time $\chi$'s is

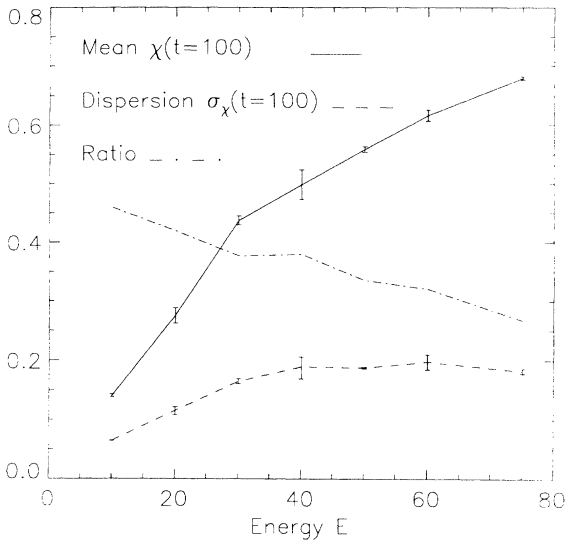


FIG. 12. The mean $\bar{\chi}(t=100)$ (solid line), the associated dispersion $\sigma_\chi(t=100)$ (dashed line), and the ratio $r=\sigma_\chi(t=100)/\bar{\chi}(t=100)$ (dot-dashed line) for 1200 orbits at each energy $E$, corresponding to initial conditions chosen so as to sample the invariant measure.

quite broad, with a dispersion that is typically $\sim 30-40\,\%$ as large as the mean. This behavior is illustrated in Fig. 12. The curves in that figure were constructed by grouping together the total of 1200 orbits from all three ensembles at each energy $E$, and, for the composite distributions of $\chi$'s, computing a total mean and dispersion. The three curves in the figure represent $\bar{\chi}$, $\sigma_\chi$, and the ratio $r=\sigma_\chi/\bar{\chi}$, each evaluated at time $t=100$. To obtain estimates of the errors associated with these quantities, the three different ensembles at each energy were also analyzed individually. The plotted error bars for $\bar{\chi}$ and $\sigma_\chi$ reflect the computed standard deviations associated with the means and dispersions of the individual ensembles.

It should perhaps be stressed again that, from the viewpoint of nonlinear dynamics, there is nothing special about the particular value $t=100$ which has entered into all of these calculations: one can equally well compute the mean $\bar{\chi}$ for shorter or longer times. Given that the initial ensemble does not sample the invariant measure, it is hardly surprising that, at early times $t \ll 100$, $\bar{\chi}(t)$ does not coincide with the true Lyapunov exponent. However, as the ensemble converges towards the invariant measure $\bar{\chi}$ should also converge towards the true Lyapunov exponent $\chi$; and indeed, a numerical evaluation of $\bar{\chi}(t)$ shows that, for $t > 50$ or so, $\chi$ and $\bar{\chi}(t)$ agree quite closely [19].

Computing $\chi(t)$ for 400 orbits for a total interval $t=50$ only requires as much time as computing $\chi(t)$ for a single orbit over a period $t=2\times 10^4$ or for a pair of orbits for $t=10^4$. However, at least for this model potential, such a short times computation provides a more reliable estimate overall of the true Lyapunov exponent than what would be derived by tracking one or two orbits for a time $t=10^4$. In this sense, one can argue that the short times estimate of $\chi$, in addition to being more physical when considering systems on relatively short time scales, is computationally useful as well, in that it serves as a more efficient algorithm than the usual procedure.

One final remaining point is that the short times procedure provides substantially more information than does the standard long time estimate of $\chi$. Evaluating $\chi(t)$ for all of the orbits in some ensemble yields a *distribution of instability time scales,* which provides a potentially useful characterization of the degree to which different orbits are unstable on different time scales. This fact has already been exploited by various authors for other systems in a variety of different contexts [15,16]. Applications of this approach to the transient dynamics of ensembles of orbits in galactic potentials are discussed elsewhere [19].

## ACKNOWLEDGMENTS

The authors acknowledge useful discussions with Robert Abernathy, Salman Habib, and Paul Channell. H.E.K. also acknowledges useful discussions with Dev Thirumalai and Edward Ott. H.E.K. was supported in part by the National Science Foundation Grant No. PHY92-03333. M.E.M. was supported in part by the University of Florida. Some of the calculations reported herein were effected using computer time made available through the Research Computing Initiative at the Northeast Regional Data Center (Florida) by arrangement with IBM.

[1] Cf. C. W. Allen, *Astrophysical Quantities,* 3rd ed. (Athlone, London, 1973).

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

[3] Cf. S. Udry and D. Pfenniger, Astron. Astrophys. **198**, 135 (1988).

[4] Cf. B. A. Dubrovin, I. M. Krichever, and S. P. Novikov, in *Dynamical Systems IV: Encyclopedia of Mathematical Sciences,* edited by V. I. Arnol'd and S. P. Novikov (Springer-Verlag, Berlin, 1990).

[5] Cf. A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics* (Springer-Verlag, Berlin, 1992).

[6] Since the growth of any initial error is exponential, even very small errors in the specification of initial data are amplified extremely quickly. The time required for a small error to grow into a macroscopic deviation in the orbit depends only logarithmically on the size of the initial error.

[7] D. V. Anosov, Proc. Steklov Inst. Math. **90**, 1 (1967); R. Bowen, J. Differ. Eq. **18**, 333 (1975); C. Grebogi, S. M. Hammel, J. A. Yorke, and T. Sauer, Phys. Rev. Lett. **65**, 1527 (1990); T. Sauer and J. A. Yorke, Nonlinearity **4**, 961 (1991).

[8] H. E. Kandrup, Phys. Lett. A **140**, 97 (1989); Physica A **169**, 73 (1990).

[9] P. Cipriani and G. Pucacco, Astron. Astrophys. (to be published).

[10] Cf. R. H. Miller, Astrophys. J. **140**, 250 (1964); H. E. Kandrup and H. Smith, *ibid.* **374**, 255 (1991); **386**, 635 (1992); H. E. Kandrup, H. Smith, and D. E. Willmes, *ibid.*

**399**, 627 (1992); H. E. Kandrup, M. E. Mahon, and H. Smith, *ibid.* (to be published); J. Goodman, D. Heggie, and P. Hut, *ibid.* **415**, 715 (1993).

[11] Cf. numerous references cited in J. Binney and S. Tremaine, *Galactic Dynamics* (Princeton University Press, Princeton, NJ, 1987).

[12] M. Toda, J. Phys. Soc. Jpn. **22**, 431 (1967); **23**, 501 (1967).

[13] G. Contopoulos and C. Polymilis, Physica D **24**, 328 (1987).

[14] G. D. Birkhoff, Proc. Natl. Acad. Sci. USA **17**, 656.

[15] Cf. P. Grassberger, R. Badii, and A. Poloti, J. Stat. Phys. **51**, 135 (1988).

[16] Cf. M. A. Sepúlveda, R. Badii, and E. Pollak, Phys. Rev. Lett. **63**, 1226 (1989); F. Varosi, T. M. Antonsen, and E. Ott, Phys. Fluids A **3**, 1017 (1991); J. M. Finn, J. D. Hanson, I. Kan, and E. Ott, Phys. Fluids B **3**, 1250 (1991).

[17] V. I. Oseldec, Tr. Mosk. Ova. Ispyt. Prir. **19**, 179 (1968) [Trans. Mosc. Math. Soc. **19**, 197 (1968)].

[18] R. S. Shirts and W. P. Reinhart, J. Chem. Phys. **77**, 5204 (1982); Y.-T. Lau, J. M. Finn, and E. Ott, Phys. Rev. Lett. **66**, 978 (1991); G. Contopoulos, H. E. Kandrup, and D. Kaufmann, Physica D **64**, 310 (1993).

[19] H. E. Kandrup and M. E. Mahon, Astron. Astrophys. (to be published).

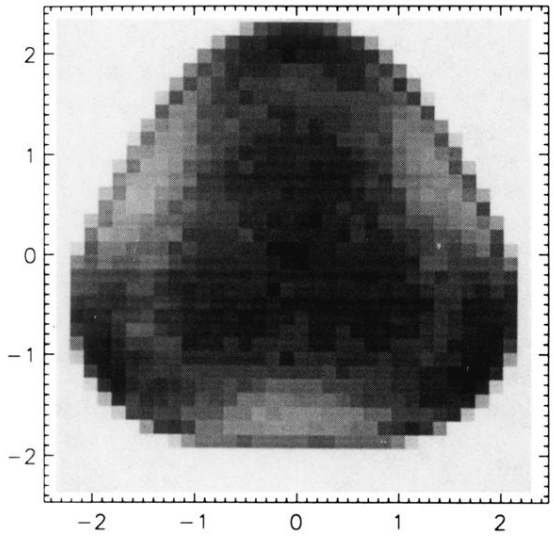


FIG. 8. A gray-scale plot of the projected invariant distribution $f_{\text{inv}}(x,y)$ for stochastic orbits with $E=30$.

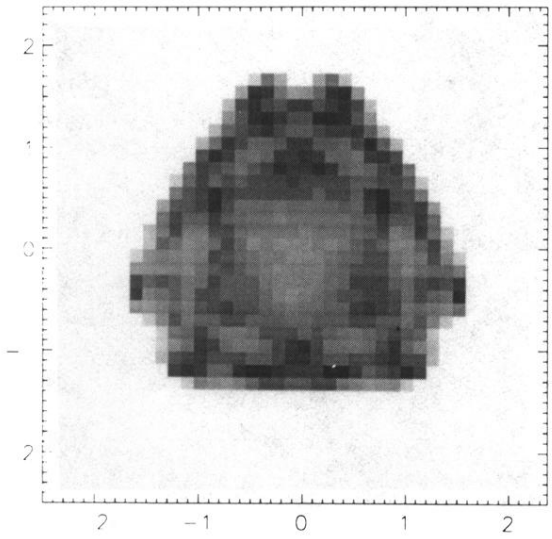


FIG. 9. A gray-scale plot of the projected invariant distribution $f_{\text{inv}}(x,y)$ for stochastic orbits with $E=10$.