2005

Coarse-grained entropy decrease and phase-space focusing in Hamiltonian dynamics

Arjendu K. Pattanayak
Carleton College

Daniel W. C. Brooks
Carleton College

Anton de la Fuente
Carleton College

Lawrence Uricchio
Carleton College

Edward Holby
Carleton College

See next page for additional authors

Coarse-grained entropy decrease and phase-space focusing in Hamiltonian dynamics

Arjendu K. Pattanayak, Daniel W. C. Brooks, Anton de la Fuente, Lawrence Uricchio, Edward Holby, Daniel Krawisz, and Jorge I. Silva
Department of Physics, Carleton College, Northfield, Minnesota 55057, USA
(Received 31 January 2005; published 15 July 2005)

We analyze the behavior of the coarse-grained entropy for classical probabilities in nonlinear Hamiltonians. We focus on the result that if the trajectory dynamics are integrable, the probability ensemble shows transient increases in the coherence, corresponding to an increase in localization of the ensemble and hence the phase-space density of the ensemble. We discuss the connection of these dynamics to the problem of cooling in atomic ensembles. We show how these dynamics can be understood in terms of the behavior of individual trajectories, allowing us to manipulate ensembles to create "cold" dense final ensembles. We illustrate these results with an analysis of the behavior of particular nonlinear integrable systems, including discussions of the spin-echo effect and the seeming violation of Liouville's theorem.

DOI: 10.1103/PhysRevA.72.013406 PACS number(s): 32.80.Pj, 03.65.Sq, 02.50.−r, 05.20.Gg

I. INTRODUCTION

The evolution of probability distributions in nonlinear Hamiltonian systems is fundamental to nonequilibrium statistical mechanics [1]. The rich behavior that results when the classical point dynamics are mapped onto the quantum and classical probability dynamics continues to provide intriguing problems for study. Recent work [2,3] has used a coarse-grained analysis of such Hamiltonian systems to study the interesting phenomenon of coherence enhancement for an ensemble. This considered the behavior of the entropy for a coarse-grained classical distribution evolving in a nonlinear integrable Hamiltonian. A coarse-grained distribution (defined more carefully below) results from a Liouville probability distribution being smoothed with a function (for example, a Gaussian, possibly representing the measurement resolution limits in each phase-space dimension) at each phase-space point. It may also be constructed by considering the behavior of many individual trajectories, and by counting the number in small but finite-sized bins in phase-space. The analysis showed that the entropy of the coarse-grained distributions could actually be significantly *decreased*.

The entropy of the standard unsmoothed distribution does not change for Hamiltonian systems, as expected from Liouville's theorem. However, the coarse-grained entropy and density can indeed change and in particular be improved via Hamiltonian dynamics. In that coarse-graining seems artificial, the question that must immediately be answered is: What is the physical meaning of the coarse-grained results; are they real, in short? We argue below in some detail that coarse-grained dynamics are physically meaningful for various reasons: (i) They arise naturally as a result of a physical lack of infinite resolution in phase-space. (ii) Coarse-graining also applies whenever any real classical ensemble (formed from a finite set of particles) is considered. (iii) Most convincingly, this change in the coarse-grained quantities can be mapped to experimental signals, as also discussed further below. In particular, this entropy decrease corresponds to an increase in the coherence of the distribution, i.e., to focusing in phase-space.

There are potentially exciting applications for this interesting result of coarse-grained entropy decrease and phase-space focusing. For some systems, as the localization area in phase-space $\to\hbar$, quantum effects are expected to be significantly enhanced. When the quantal and classical dynamics follow each other, this focusing serves as a basis for algorithms to generate sharply localized wave packets with nonlinear techniques. This specific application has begun to be explored: Excited electrons in Rydberg atoms were studied [2] to predict that an initial equilibrium ensemble could be focused tightly in phase-space. This has since been experimentally verified [3] using an ensemble of potassium atoms. Another application is to cooling and enhancing phase-space density for an ensemble of particles, a problem at the forefront of current experimental physics [4]. Typical approaches to focusing rely on dissipative techniques, which means that they are difficult to apply to systems without accessible internal degrees of freedom. As such, Hamiltonian methods would greatly expand our ability to cool arbitrary systems. The cooling application has not yet been explored and is part of the motivation for this paper.

Below, we present further studies of coarse-grained entropy (CGE) dynamics, showing in particular how CGE oscillations relate to phase-space dynamics and hence how this can be used in specific physical applications. Our goal here is primarily to explore the physical basis of this behavior. We first lay out a relatively formal analysis of the connection between coarse-graining, entropy dynamics, and the behavior of trajectories, including some general protocols. However, these abstract results do not provide rules for specific choices of Hamiltonians, parameters, and initial conditions. We then translate the broader concepts to more intuitive ideas about the behavior of trajectories, by looking at some applications, discussing in particular the connection to the spin-echo effect, for example. We also explore the connection to recent information-theoretic perspectives on cooling of atomic ensembles. We then consider some results from specific Hamiltonians and show that at least two different aspects of Hamiltonian trajectory dynamics can lead to coherence enhancement. We discuss the subtle ways in which fundamental results such as Liouville's theorem apply in these situations, and conclude by discussing specific applications, including the cooling of ensembles.

II. TRAJECTORIES AND COARSE-GRAINED ENTROPY DYNAMICS

We derive here the connection between CGE dynamics and trajectory dynamics; the analysis follows previous work [2,5] but also makes this connection explicit. The crucial part of this section are the results showing entropy oscillations in general integrable Hamiltonians, Eq. ((27) in particular.

Consider a distribution $\rho(\vec{x},t)$ with the phase-space variables $\vec{x}(t)=(\vec{p},\vec{q})$ denoting both the momentum and configuration variables. This distribution may represent a cloud of noninteracting particles, multiple realizations of single-particle dynamics, or our knowledge of a single particle's dynamics with associated uncertainty. In some limit, these quantum dynamics of all these situations are expected to follow the behavior of this distribution, and many of our results apply to quantum-mechanical systems as well—this is discussed briefly later.

The Liouville equation

$$\frac{\partial \rho}{\partial t} = \{H,\rho\} \tag{1}$$

applies to the general classical situation, where $\{A,B\}$ is the Poisson bracket for the phase-space functions A, B. We now show that the coherence of such distributions oscillates in complete generality for nonlinear integrable Hamiltonians.

To start, we need a measure for the coherence. The coherence arguably represents the extent to which members of the ensemble have similar properties, or alternatively, the degree to which we have certainty about the location of a single particle in phase-space. Given an information-theoretic perspective, it is intuitive that a reasonable measure of the degree of statistical coherence of $\rho(\vec{x},t)$ is provided by some form of entropy for the distribution. Either the Gibbs entropy $-\mathrm{Tr}[\rho \ln(\rho)]$ or the analytically more useful Renyi entropy [6] $-\ln[\mathrm{Tr}(\rho^2)]$ may be used, where Tr denotes the trace or integral over all phase space variables. However, as is known, since the Liouville equation can also be written as $d\rho[\vec{x}(t),t]/dt=0$, all functions of ρ, including the entropy, are constant in time for systems evolving under the Liouville equation. This property of the fine-grained distribution (that is, a distribution that is *not* coarse-grained in any way) is singular. That is, the behavior of a distribution that is coarse-grained by any amount, no matter how small, is distinctly different from that of the fine-grained distribution. This applies to both classical and quantal distributions. The coarse-graining in these cases could be static (due to resolution/measurement coarse-graining). Of course, in any realistic situation the system of interest is coupled, albeit weakly, to other systems. Even a zero-temperature quantum-mechanical system is coupled to the vacuum. This is typically represented by taking the deterministic equations of motion and adding noise with some specified spectrum to represent the ignored degrees of freedom. Again, the behavior of the system for vanishing coupling to the environment is of interest. This is usually termed dynamical coarse-graining. In what follows, we assume that we are near the zero limit for both coarse-grainings, but also that the static coarse-graining masks dynamical coarse-graining on the time scales of interest. Finally, for classical mechanics, a continuous ρ does not represent a real classical ensemble of point particles as accurately as the coarse-grained construction defined below. Thus, to usefully and accurately understand the behavior of distributions, a small but finite coarse-graining must be used. We show more carefully below that this coarse-graining is particularly useful in distinguishing a coherent distribution from one that is more spread out.

One way to coarse-grain a distribution therefore is to divide phase-space into equally spaced cells defined by vectors of length $\delta\vec{x}$ centered at phase-space points $\vec{x}_i$ where $i=1, 2, \ldots$ labels the cells in phase-space. These cells have hypervolume

$$\delta A = \int_0^{\delta\vec{x}} d\vec{x}. \tag{2}$$

The values of the densities associated with each of these phase-space points correspond to the average value of the density within a given cell [7],

$$\rho_i(t) = \frac{1}{\delta A}\int_{\vec{x}_i-\delta\vec{x}/2}^{\vec{x}_i+\delta\vec{x}/2} d\vec{x}\rho(p,q,t). \tag{3}$$

The coarse-grained Renyi entropy can then be defined in the usual way [7] as

$$S_c(t) = -\ln\left[\delta A\sum_{i=1}^{\infty}\rho_i^2(t)\right]. \tag{4}$$

Our interest therefore is in the dynamics of S_c.

Instead of solving for the behavior of the continuum object ρ and then coarse-graining to find ρ_i, we computationally follow a method termed the classical trajectory Monte Carlo method (see Ref. [2], for example). A large but finite ensemble of N_{traj} phase-space points are generated, distributed according to a chosen initial density. Each point is then evolved forward using the Hamiltonian equations of motion. The coarse-grained density is then numerically calculated as

$$\rho_i^{(C)} = \frac{1}{\delta A}\frac{N_i}{N_{\mathrm{traj}}}, \tag{5}$$

where N_i is the fraction of the trajectories lying in the cell centered at $\vec{x}_i$, and the CGE calculated from this as above. This quantity then goes over identically to the coarse-grained density obtained as in Eq. (3) from the continuum field in the limit $\lim_{N_{\mathrm{traj}}\to\infty}$.

We now analyze the dynamics of $S_c(t)$ as a function of time for general nonlinear Hamiltonians. This is best understood analytically by using another representation of coarse-graining where we locally smooth the distribution $\rho(\vec{x},t)$ by a Gaussian of width $\delta_{\vec{x}}$. An important point to note here is that the predictions for coherence enhancement proceed from this kind of local coarse-graining, as opposed to requiring trajectory binning, which means that it applies in principle to quantum dynamics as well. The coarse-grained Renyi entropy can be written for the Gaussian smoothing case as [8]

$$S_c = -\ln \mathrm{Tr}\left[\left(\exp\left(-\delta\sum_i \frac{\partial^2}{\partial x_i^2}\right)\rho(\vec{x})\right)^2\right]. \quad (6)$$

It is useful to consider these quantities in the Fourier representation of ρ, also called its characteristic function. In terms of the Fourier transform $\tilde{\rho}(\vec{k})$ of ρ, the Renyi entropy becomes

$$S_c = -\ln\{\mathrm{Tr}[\exp(-2\delta k^2)|\tilde{\rho}(\vec{k})|^2]\}, \quad (7)$$

where the trace now runs over k values. In the limit $\delta \to 0$, S_c can be written

$$S_c = -\ln(\mathrm{Tr}[\tilde{\rho}^2] - 2\delta\,\mathrm{Tr}[k^2|\tilde{\rho}|^2]) + \mathcal{O}(\delta^2) \quad (8)$$

$$= -\ln\{(1-2\delta\chi^2)\mathrm{Tr}[\tilde{\rho}^2]\} + \mathcal{O}(\delta^2). \quad (9)$$

Here we have defined $\chi^2 \equiv \mathrm{Tr}(k^2|\tilde{\rho}|^2)/\mathrm{Tr}(|\tilde{\rho}|^2)$ which is the second moment of $\vec{k}$ for $\tilde{\rho}$. The quantity χ^2 can be understood as measuring the degree of fine-scale "structure" in a distribution. That is, as a distribution needs more Fourier components of higher k in order to be accurately specified, χ^2 increases. Notice that the terms being neglected of higher order in δ are related to k^4, k^6, etc., indicating that S_c as defined ignores higher-order information in the distribution, as appropriate.

The entropy dynamics depend on $\chi^2(t)$. This quantity has been analyzed in some detail previously [9] and has been shown to be useful in various contexts, arising from being the measure of the sensitivity of a distribution to coarse-graining. For example, an unusual recent result [5,10] is that some persistent patterns [11,12] that arise in the context of chaotic two-dimensional fluid dynamical systems are characterized by an approximately constant χ^2, which can be interpreted as resulting from a balance between chaos and noise in the system. χ^2 has also proven useful in understanding quantum-classical correspondence in chaotic systems [13,14]. One reason χ^2 is useful is that the classical trajectory dynamics can be mapped directly onto the behavior of χ^2, as we now show.

Consider a possibly time-dependent Hamiltonian $H(\vec{x},t)$ where $\vec{x} \Leftrightarrow (\vec{p},\vec{q})$. The equations of motion for $\vec{x}(t)$ are just Hamilton's equations, written here as

$$\dot{\vec{x}} = \vec{v}(\vec{x},t) \quad (10)$$

with individual components $\dot{x}_\alpha = v_\alpha(\vec{x})$. The stability of the dynamics is characterized by considering the behavior of linearized deviations from the trajectories; thus, substituting $\vec{x} + \vec{\varsigma}$ in Eq. (10), we get

$$\frac{d\vec{\varsigma}(t)}{dt} = \mathcal{J}\vec{\varsigma}(t), \quad (11)$$

where the elements of the matrix $\mathcal{J}_{\alpha\beta} = \partial v_\alpha / \partial x_\beta$. The eigenvalues of $\mathcal{J}$ in general vary along the trajectory $\vec{x}(t)$. In particular, we write the dynamics of $\vec{\varsigma}$ as

$$\vec{\varsigma}(t) = \mathcal{M}\vec{\varsigma}(0), \quad (12)$$

where $\mathcal{M}$ is given by [15] the time-ordered series

$$\mathcal{M}(t,\vec{x}(0)) = \mathcal{T}\exp\int_0^t \mathcal{J}(\tau)d\tau \quad (13)$$

and $\mathcal{T}$ is the time-ordering operator. A real symmetric matrix $\mathcal{M}^T \cdot \mathcal{M}$ (the transpose being denoted by $\mathcal{M}^T$) can now be defined and diagonalized [15] as

$$M^T \cdot M = \sum_k u_k(t,\vec{x}(0))\sigma_k(t,\vec{x}(0))u_k^T(t,\vec{x}(0)). \quad (14)$$

The u_k constitute a local orthonormal tangent space for the flow. This basis set of unit vectors is well defined at every point of the trajectory $\vec{x}(t)$, although it changes orientation at each point [note that we use Latin indices (k,l) to distinguish vector components in this basis]. In this basis, components of the matrix $\mathcal{M}^T \cdot \mathcal{M}$ can be written as $\sigma_k(t,\vec{x}(0))\delta_{k,l}$ and also

$$\sum_k \varsigma_k^2(t) = \sum_k \sigma_k(t,\vec{x}(0))\varsigma_k^2(0). \quad (15)$$

This holds independent of our choice of $\varsigma_k(0)$ and hence

$$\varsigma_k^2(t) = \sigma_k(t,\vec{x}(0))\varsigma_k^2(0) \quad (16)$$

for each individual component. For chaotic systems the eigenvalues of $\mathcal{J}$, appropriately averaged [15], are termed the Lyapunov exponents, and are the time-scale for divergence of solutions.

We now apply this formalism to ρ. The Liouville equation written in terms of the phase-space velocities $\vec{v}$ is

$$\frac{d\rho}{dt} = 0 = \frac{\partial\rho}{\partial t} + \vec{v}\cdot\vec{\nabla}\rho, \quad (17)$$

where the first equality again represents the incompressibility of ρ and the gradient operator $\vec{\nabla} = \Sigma_\alpha \hat{\alpha}\nabla_\alpha$ is defined in terms of all phase-space variables. Applying $\vec{\nabla}$ to the second equality, we get that

$$\sum_\alpha \hat{\alpha}\left[\frac{\partial}{\partial t}\nabla_\alpha\rho + \sum_\beta v_\beta \nabla_\alpha\nabla_\beta\rho + \sum_\beta \nabla_\alpha v_\beta \nabla_\beta\rho\right] = 0. \quad (18)$$

The first term and the sum over β of the second term (with the order of differentiation interchanged) is just the total derivative of $\nabla_\alpha\rho$ and the final term can be written as the sum over the product $\mathcal{J}_{\beta\alpha}\nabla_\beta\rho$. Thus $\vec{\nabla}\rho$ satisfies

$$\frac{d}{dt}\vec{\nabla}\rho(t) = -\mathcal{J}^T\vec{\nabla}\rho(t), \quad (19)$$

which is also the adjoint equation of variations. The symmetry of Eqs. (19) and (11) is intuitive: For an incompressible flow, the field ρ cannot be created or destroyed along trajectories, and hence the probability density profile sharpens (flattens) along the direction in which trajectories move closer (further). Equation (19) can be used to generate a solution for $\nabla\rho$ just as in Eq. (12). The important point is that for Hamiltonian systems, the eigenvalues σ_k that govern the behavior of the tangent-space vectors $\vec{\varsigma}$ also govern the behavior of $\vec{\nabla}$.

The critical step in connecting this detailed microscopic behavior to the dynamics of the entropy is to realize that an

average of the gradient over the entire distribution is related to χ^2 [and hence to S_c, through Eq. (9)]. In particular, $\chi^2 = \Sigma_j \chi_j^2$ where

$$\chi_j^2 \equiv \frac{\int d\vec{x} |\nabla_j \rho(\vec{x}(t))|^2}{4\pi^2 \int d\vec{x} \rho^2(x(t))} \tag{20}$$

$$= \frac{-\int d\vec{x} \rho \nabla_j^2 \rho(\vec{x}(t))}{4\pi^2 \int d\vec{x} \rho^2(x(t))} \tag{21}$$

$$= \frac{\int d\vec{k} k_j^2 |\tilde{\rho}(\vec{k})|^2}{\int d\vec{k} |\tilde{\rho}(\vec{k})|^2}, \tag{22}$$

where Eq. (21) arises from integrating by parts and the lack of contribution from boundary terms. Equation (22) follows from the definition of the Fourier transform. Finally, we can see that the equivalent of the relationship Eq. (16) for $\nabla_i \rho$ implies that

$$\chi_i^2(t) = \frac{\int d\vec{x} |\nabla_i \rho(\vec{x}(t))|^2}{4\pi^2 \int d\vec{x}\, \rho^2(\vec{x}(0))} \tag{23}$$

$$= \frac{\int d\vec{x} |\nabla_i \rho(\vec{x}(0))|^2 \sigma_i(t, \vec{x}(0))}{4\pi^2 \int d\vec{x}\, \rho^2(\vec{x}(0))} \tag{24}$$

so that the dynamics of χ^2 depend on an average over the exponentiated eigenvalues of the "stability" matrix $\mathcal{J}$, with an unusual weighting given by the initial local gradient squared.

For chaotic systems, the eigenvalues of $\mathcal{J}$ are real with nonzero averages, at least one of which is positive, and hence it can be shown that χ^2 increases exponentially rapidly such that

$$\lim_{t \to \infty} \frac{1}{t} \ln(\chi) = \Lambda_{\max}, \tag{25}$$

where $\Lambda_{\max}$ is the so-called largest Lyapunov exponent. For integrable (regular, nonchaotic) dynamics, on the other hand, the eigenvalues of $\mathcal{J}$ are imaginary, so that χ^2 can be written as a sum over periodic functions with different frequencies. In general, this is a sum with a continuous spectrum of frequencies, and hence χ^2 executes bounded dephasing oscillations. This shows clear oscillations for short times, with a slow increase for longer times.

We now use this to deduce how nonlinear Hamiltonian dynamics manipulate or change the encoded information in the distribution. First, note that since χ^2 measures the amount of structure in the distribution, it effectively measures whether the information encoded in ρ is resident at lower or higher resolution or equivalently at lower or higher degrees of correlations within members of the ensembles. The physical loss of information about higher correlations due to environmental perturbations or experimental lack of resolution corresponds analytically to an increase in entropy due to coarse-graining. This is precisely what happens as χ^2 increases, as can be seen from Eq. (7) and subsequent approximations. Explicitly, for systems not coupled to the environment, for short time scales,

$$\dot{S} \propto \dot{\chi}^2. \tag{26}$$

For nonlinear chaotic dynamics, χ^2 increases essentially monotonically [16], so that information moves unidirectionally and increase of the CGE. On the other hand, for nonlinear *integrable* dynamics the oscillating behavior of χ^2 noted above means that in full generality the information in the distribution moves back and forth between the lower and higher degrees of correlation of the distribution as a result of the natural Hamiltonian dynamics. Thus, when χ^2 is large, in a coarse-grained version of such a system we have in principle lost some information about the system. However, the deterministic dynamics dictates that χ^2 will then oscillate, or that the information will flow back into the system from the finest scales. When it does so, the state decreases its χ^2 and hence has lower CGE, and would be more coherent or localized.

This result is the central basis of the phenomenon we are exploring of coherence enhancement, establishing that in general integrable nonlinear Hamiltonian dynamics, the coherence of ensembles can be enhanced. Mathematically, if χ remains small, as is true in the short-time limit for integrable systems, a final good approximation yields

$$S_c(t) = C + 2\delta\chi^2(t) + \mathcal{O}(\delta^2), \tag{27}$$

where C is some constant dependent upon the initial distribution. Since S_c is oscillating, it can in general be *less* than its initial value at certain times during the evolution as the distribution relaxes towards its long-time limit of increased entropy. Thus transient increases in coherence as measured by the entropy of the coarse-grained distribution are a completely general phenomenon in nonlinear Hamiltonians.

Several points should be noted:

(i) This oscillation of χ^2 generally only holds for short times. As is clear below, by "short times" we mean that the oscillations occur over a few characteristic time periods of the natural periodic orbits of the trajectories in the distribution. The generic behavior of χ^2 is dephasing oscillations, or sum over oscillations with a continuum of frequencies, and as such the asymptotic behavior is that χ^2 grows slowly. Specifically, unlike chaos, this growth is not exponentially rapid.

(ii) The value of χ^2 nominally does not affect the fine grained entropy of the distribution (that is, the entropy computed with no added coarse-graining). However, some sort of coarse-graining always exists, including, for example,

particle-particle collisions in atomic ensembles, and the CGE oscillation therefore is a real result. This means that as a general result, the entropy of a distribution in a nonlinear integrable Hamiltonian system will oscillate for short times.

(iii) δ is not a property of the "system" in that it represents an "external" resolution or measurement, so that the system does not know about it. It is therefore a good question to ask how δ affects the size of the oscillation of the coherence. By examining the definition of the coherence we can see that all δ does is set the zero and an entirely arbitrary scale for the oscillation. As is appropriate, physical measures corresponding to experimentally observable signals must be used to set the scale of the oscillation; these are idiosyncratic for each application.

(iv) Similarly, we note that this oscillation is mathematically entirely generic—that is, it holds in principle for arbitrary choices of initial conditions and integrable nonlinear Hamiltonians. Physically, however, it is clear that initial conditions critically matter to the size of the visible effect. A near-equilibrium (or almost completely dephased) distribution for a given Hamiltonian will not show significant physically relevant enhancement of coherence. Also, that the coherence generically oscillates for a distribution does not guarantee that the future state of the distribution will necessarily improve on the initial state. Equivalently, the time scales of oscillations might not be physically useful. As such, careful analysis is in general important such that this result can be constructively applied.

A general protocol using this result for coherence enhancement suggests itself: Let the initial distribution be an equilibrium distribution for some nonlinear Hamiltonian. At this point, either (a) apply an abrupt perturbation or (b) change the Hamiltonian. The ensemble is now a nonequilibrium distribution and will undergo some dynamics on the way to equilibrium. At times long before it has relaxed to equilibrium, there will be situations where it is transiently more focused than it was initially. At this point, if appropriate, the coherence of the ensemble could be used directly. Alternatively, it is possible to apply an appropriate Hamiltonian at this point such that the distribution is then "trapped" in phase-space, and is in a state of greater coherence than it was initially.

Having derived the formal relationship between classical trajectory dynamics and the CGE of an ensemble, and presented a broad protocol for exploiting this, we now turn to a more careful analysis, specifically to understand the physical signatures of changes in the CGE.

III. COARSE-GRAINED ENTROPY AND PHYSICAL MEANING

We now consider how the coarse-grained entropy translates to physical signatures, particularly a broader intuitive discussion of the issues to help enhance our physical understanding. We maintain this intuitive perspective in the rest of the paper.

It is useful to think about these issues visually: The difference between the fine-grained and CGE (or density) is roughly the difference between the (hyper)volume and surface (hyper)area of the phase-space distribution: The volume remains constant but the bounding surface can change fairly dramatically in size. This is easiest visualized in a two-dimensional phase-space, where this translates to a constant area with changing perimeter.

Now consider a hypothetical distribution in phase-space, and the various possible shapes into which it can be manipulated while retaining its area. The coarse-grained entropy results indicate that it is possible that the area changes dramatically as the distribution changes shape while undergoing Hamiltonian dynamics, which is perhaps not too surprising. However, what is interesting is that this difference in shape has physical and experimental meaning. This can be understood as follows: Suppose that an experimental signal is associated with phase-space location, and define this signal as $F(q,p)$ for what follows. If F changes reasonably rapidly as a function of phase-space distance; that is, $F(q_1,p_1)$ is significantly different from $F(q_2,p_2)$ for $(q_1-q_2)^2+(p_1-p_2)^2$ small, then the shape of the distribution is going to matter a great deal in determining the nature and coherence (the range in F obtained) of the signal. The subjective terms "small" and "significantly different" scale as required by the experimental signal itself, as well as by the resolution limits with which we are concerned.

A. CGE increase and chaotic dynamics

To clarify some of the paradoxes of constant entropy, consider the behavior of a chaotic system. In this situation, if we start with a sharply localized distribution and allow it to evolve, the fine-grained entropy does not change. The lack of change of the fine-grained entropy in this case creates a conceptual paradox. It seems to imply that although the chaotic dynamics of an individual trajectory causes us to lose information about it, this information is still maintained somehow when considering the entire distribution. The information loss-gain reconciliation is simple enough [9,19]: The same finite resolution for trajectories that results in error has to be included in the analysis of distributions, at the very least to see if predictions in the limit of zero infintesimal coarse-graining agrees with the fine-grained behavior. That is, information in chaotic systems only flows from the lower-order correlations to the higher-order correlations. This one-way transfer of information about the distribution from the grosser scales to the finest scales proceeds exponentially rapidly, reflected in the exponentially rapid growth of χ^2. As such, for nonlinear *chaotic* dynamics, the natural dynamics causes rapid loss of information during the initial stages. The CGE therefore changes dramatically in a chaotic system, decreasing exponentially rapidly initially, followed by a later linear decay with the time scales set by the Lyapunov exponent of the system [5,13,20,21]. In particular, the critical point is that the chaotic dynamics spreads the distribution in a complicated way across phase-space while still maintaining the same area. In this case again, the physically relevant quantity turns out to be the coarse-grained description, and the fine-grained description is a singular limit.

B. CGE dynamics during spin echo

Consider now how this applies to spin-echo dynamics [17]. The initial condition is an ensemble of spins, all ini-

tially aligned. In a simple quantitative model, each spin is a dipole of moment $\vec{m}$ fixed at its center but free to rotate in the presence of a constant external magnetic field $\vec{B}$. The equation of motion for each dipole is

$$\frac{d\,\vec{m}(t)}{dt} = g\,\vec{m}(t)X\vec{B}, \tag{28}$$

where g is the gyromagnetic ratio. When released from rest the dipole will precess at a constant angle to **B**. In particular, if $\vec{B}$ is chosen to be the negative z axis and if initially $\vec{m}$ is in the x-y plane, the vector remains in the x-y plane and is given by

$$\vec{m}(t) = (m\cos(\omega t), m\sin(\omega t)), \tag{29}$$

where $\omega = g|B|$, such that the tip of the vector traces out a circle in the x-y plane. This plane is then the phase-space for this particular dynamical system. Also, the experimental signal here relates to the particular direction in which the dipole spin is pointing, that is, the location on a phase-space circle for the spin.

In considering the ensemble's behavior, note that in general the magnetic field is inhomogeneous. For the moment, consider spatial inhomogeneities only, such that each spin is evolving in a slightly different field, and ignore any temporal fluctuations in the field, including any effect of spin-spin interactions. This system then maps precisely to an ensemble of particles distributed on different orbits in a nonlinear integrable Hamiltonian. In particular, each spin (particle) moves with a different frequency depending on the local field.

Without loss of generality, we may visualize this by assuming that in phase-space, increasing radial distance from the phase-space origin corresponds to increasing frequency. The initial ensemble with the spins almost perfectly aligned then corresponds to an initial phase-space distribution localized very tightly in angle, but with a certain spread in radius. As the system evolves, the variation in the various frequencies of rotation for the spins means that each spin moves at a different rate along its phase-space circle. That is, some orbits move faster than others so that the various periodic solutions dephase, and the experimental signal (the magnetization density) degrades. The faster spins are "ahead" in their orbits, which corresponds to the original tightly localized distribution being stretched out into a spiral-like tendril in the long run. Note that, as required by Liouville's theorem, this tendril maintains the same area as the tightly localized initial condition.

The interesting part of the spin-echo system is, of course, that this experimental degradation of the signal can be reversed. This can be done in one of two different ways. The first is if the magnetic field is exactly reversed at some time $t=T$. This would mean that that each orbit would now move "backwards" and retrace its steps, and in particular that the faster spins would be "behind," such that the distribution would congeal back at $t=2T$ into the same tightly localized distribution as before. An alternative is to apply a magnetic field in the plane for a short time such that every particle is rotated around the x axis back onto the x-y plane, say. The faster spins are behind again, and the same tightly localized distribution obtains again in the future. The key point here is that during the "echo" dynamics the different frequencies of evolution work to align an incoherent ensemble.

The fine-grained entropy in this case would remain constant throughout the evolution; however, the CGE increases during the dephasing, and then decreases during the echo dynamics. This also happens to the experimental signal, whence the CGE is mapped to a physically relevant quantity in this situation.

We clarify an interesting and potentially somewhat confusing point: Under ideal conditions, the distribution at $t=2T$ after spin-echo is identical to the distribution at $t=0$, and so are the values of the CGE at these two times; the fine-grained entropy is constant, again, for the entire dynamics. However, under realistic conditions, there are local time-dependent fluctuations of the magnetic field such that the original distribution is not exactly recovered by the echo. This corresponds to a decay in the fine-grained entropy [18] as well as the CGE and is a different effect from the one we are currently considering. However, the rate at which the entropy decays can also be mapped onto the classical dynamics of the system for chaotic systems.

C. Rydberg atom ensembles

An experimental implementation of the protocol suggested above for coherence enhancement has been achieved for an ensemble consisting of the outermost electron in a set of Rydberg atoms [3]. A remarkably accurate one-dimensional representation of this system is given in terms of the electron's position q, momentum p, and angular momentum Λ, by the model time-dependent Hamiltonian

$$H = \frac{p^2}{2} - \frac{1}{q} + \frac{\Lambda^2}{2q^2} + \sum_{n=0}^{N} q\Delta p\ \delta(t-nT), \tag{30}$$

in scaled units. An external "delta kicking" is also shown in the delta-function term: This is the effect of a half-cycle-pulse (HCP) of a strong time-dependent electric field that can be applied for very short times to this ensemble as needed. In the limit that the pulse time is much smaller than the natural characteristic period of the orbits and essentially occupies an instant in time, this is a *kick*, that can be applied once ($N=1$ in the sum above) or periodically with spacing T as needed. Note that the application of a kick derives from a time-dependent Hamiltonian system. This is a nonlinear system with mixed chaotic and regular phase-space when kicked, and it is nonlinear integrable without the kicking. This phase-space structure was exploited to manipulate the phase-space distribution of the ensemble of Rydberg atoms and increase its statistical coherence. The ensemble in this case consisted of many parallel assays of the same system.

The enhanced coherence behavior was studied theoretically [2] in the Rydberg atom problem using the dynamics of the coarse-grained entropy. The increased coherence was then mapped onto an experimental signal, the ionization probability for the electrons. That is, it was shown that starting with an initial condition consisting of the microcanonical

distribution, and following the prescription above, the coherence of the experimental signal in response to an ionization probe could be increased dramatically. Specifically, the ensemble was manipulated such that the number of atoms that were ionized changed from a relatively flat function of the probe signal to something that was highly peaked at a specific voltage [3].

Further analysis [22] shows that the quantum results differ from the classical results as a function of the ratio $\gamma = \hbar_{\text{eff}}/|\Delta p_0|$ where the effective Planck's constant $\hbar_{\text{eff}} = 1/n$ and n is the principal quantum number for the Rydberg system, and the kick is scaled $\Delta p_0 = \Delta p/n$. As $\gamma \to 0$ the quantum results approach the classical one.

The analysis in this section confirms that the coarse-grained entropy for Hamiltonian ensembles is a physical measure with some distinct advantages in understanding dynamics. Specifically, changes in the coarse-grained entropy can be mapped onto physically and experimentally relevant concepts, and correspond to changes in the effective focusing or coherence in phase-space for the ensemble. The lack of change for the fine-grained entropy of a distribution, on the other hand, can disguise physically interesting changes in the ensemble. Arguably therefore the fine-grained entropy being constant is not the absolute constraint that it has been supposed to be for the issue of phase-space focusing in Hamiltonian dynamics. An alternative way of thinking about this is that while Liouville's theorem on phase-space volume conservation does not tell us anything about the geometry of that volume, coarse-graining analysis allows us to deduce if this volume is physically distributed in interesting ways.

D. Information and cooling

Given these abstract results about entropy dynamics, we now try to understand this in terms of its relationship to the cooling of atomic ensembles, which has been studied extensively in recent years. Since it is usually understood that Hamiltonian dynamics do not apply to the increase of phase-space density, some careful attention is paid to these ideas, and to explore the relationship to "informational" cooling.

The relationship between coherence and cooling is subtle, since coherence does not directly translate to a temperature. In the absence of equilibrium descriptions for the ensemble, the temperature of an ensemble may be thought of as approximately given by the average kinetic energy of the particles. Since the kinetic energy of a particle is $\propto p^2$, the temperature of a typical ensemble of cold atoms can then be mapped onto the width in momentum direction of a phase-space ensemble centered at $p=0$, for example. As such, it seems that decreased temperature implies increased coherence. However, since the momentum width tells us nothing about the width of the distribution in the position direction, this does not apply. Likewise, it is easy to imagine that even a tightly localized ρ could in general be centered at any value of $\langle p \rangle$, and could thus have high coherence but also high average kinetic energy. As such, a simple pair of characterizations are that cooling corresponds specifically to the increase of density in *momentum space* at $p=0$, whereas coherence corresponds to the increase of density in *phase space* at arbitrary (q,p).

Moreover, in a classic paper [23], Ketterle and Pritchard point out that the exciting developments in atomic physics arise from increasing the phase-space density (number of particles per unit phase-space volume) at the same time as decreasing the average energy of the ensemble. As such, in some sense the notion of cooling should be applied only when both of these effects apply. They also point out that when applied to a single particle (say an ion in a trap), enhanced coherence is inherently the issue of concern. That is, some particle is initially in some unknown state in a relatively large phase-space volume. If one had detailed knowledge about the particle's location in phase-space (that is, if ρ was sharply localized), applying the appropriate fields would transfer this particle to the lowest energy state. This corresponds to the assumption that a distribution with an arbitrary $\langle q,p \rangle$ can be moved without dissipation to a phase-space point $\langle p \rangle = 0$; that is, effecting the transfer of a specific amount of momentum is easier than decreasing the "width" in kinetic energy of an ensemble or the phase-space function for a single particle. It is also true in general that changing the expectation value of the position of the distribution to $\langle q \rangle = 0$ is relatively simple. Such an argument also applies to the many-particle (noninteracting) case. In this case, an appropriate field for changing $\langle p \rangle$ could be one where all the particles experience essentially the same force. One focus of Ketterle and Pritchard's paper is the statement that, given Liouville's theorem on phase-space volume conservation, Hamiltonian dynamics without feedback cannot be used to cool, or more specifically to achieve increased phase-space density or coherence. Typically, in fact, cooling is achieved by using dissipative effects from light scattering off atoms, atom-atom scattering, evaporation, etc., that is, from coupling to the internal degrees of freedom of the atom.

However, Hamiltonian techniques have been suggested for manipulating phase-space distributions so that they are cooler and denser, usually separately. The value of Hamiltonian techniques is that they can then be applied in arbitrary situations, independent of access to internal degrees of freedom for the particles in question and as such have a potentially broader application.

Fundamentally, all of these can be understood in terms of informational cooling, where we trade information about higher-order correlations of the ensemble for entropy decrease corresponding to the lower-order correlations (such as the momentum width or phase-space extent). An illustration given by Ketterle and Pritchard is as follows: If you happen to know there were no atoms near a particular wall of a container, you could move these walls in closer, thus achieving higher phase-space density. The information about the highly localized region of phase-space near the container wall certainly corresponds here to a very high-order correlation function of the ensemble.

A systematic version of this is termed "stochastic cooling" [24]. The protocol, starting with an ensemble of noninteracting particles, is as follows: (i) Measure the mean momentum of some small subset of the particles in the ensemble. (ii) Now apply the negative of the mean momentum back to the subset to move this cloud such that it is centered at the origin in momentum space. Apply this technique repeatedly to dif-

ferent subsets. To make sure that each measurement addresses a different subset of the ensemble, a good systematic step is (iii) allow the dynamics to "re-mix" the particles, before then going back to (i). The mean value of the momentum of a subset of the ensemble indeed corresponds to higher-order information for the full ensemble. What this achieves is a net movement of particles from the "tails" of the momentum distribution to the center. The net result of applying this protocol is to increase the density *and* decrease the temperature of the distribution simultaneously. In this case (as in the case of Ketterle and Pritchard's anecdotal "moving the walls" method), the agents of the information transfer between the higher-order and lower-order moments of the distribution are the experimentalists. The technical issues in the application of this algorithm (as with the moving-the-walls method) are concerned with the ability to address this subset of the ensemble so as to use the information about the higher-order moments.

E. CGE increase in delta-kick cooling

Given the technological challenges of measurement and feedback, it is tempting to think of how trapping potentials might be chosen such that the natural dynamics "automatically" effect this transfer of information, i.e., change the characteristics of the distribution in useful ways without our intervention. For example, consider "delta-kick cooling," which uses Hamiltonian dynamics for the free-particle and harmonic oscillator to manipulate the shape of the ensemble [25,26]. We illustrate a slightly modified version of this technique with no kicking. Start with a Gaussian ensemble with no $q-p$ correlation, for illustration purposes, although the method works on arbitrary initial conditions. If we draw a line along the full width at half maximum of the Gaussian, this ensemble can be pictured as an ellipse in phase-space with major and minor axes aligned with the position (along the horizontal) and momentum (along the vertical) directions. Now let ρ evolve under the free-particle Hamiltonian, that is, with no trapping potential. Particles with initial greater momentum now move faster. This shears the Gaussian ellipse out along the q axis such that one of the axes is now along a $q-p$ direction that is being rotated clockwise as a function of time, and the other axis, which must remain at 90° to this, is therefore also being rotated. More to the point, as the shearing continues the Gaussian becomes more and more like a cigar in shape, with one width growing and the other shrinking.

At some suitable point (in principle the method improves monotonically as free-particle evolution is allowed to occur for longer times, $t\rightarrow\infty$), change the evolution dynamics to that of a harmonic oscillator. That is, change the dynamics by applying the potential $V_{ho}(q)=(1/2)kq^2$ to the system. In a harmonic oscillator with suitably scaled variables, all solutions travel in circles with the same frequency. As such, the effect of applying the harmonic oscillator potential is to rotate the cigar-shaped Gaussian ellipse of the distribution. If V_{ho} is applied for the correct length of time, the distribution rotates around such that the narrow part of the cigar-shaped Gaussian is aligned with the momentum axis and the long part with the position axis. The net result is that this has lowered the momentum width (or average kinetic energy or the temperature) of the ensemble. In principle, there is no formal lower limit to the width of the distribution using this protocol and moreover, no issues of trying to address specific highly localized subensembles. This simple and direct method (sans feedback) works to decrease the momentum width of an ensemble. The nominal disadvantage of this method is that the decrease in the momentum width comes at the expense of the increase in position width.

Formally [that is, when analyzing the total distribution $\rho(q,p)$], Liouville's theorem applies to delta-kick cooling because the dynamics are entirely Hamiltonian. As such, this is formally an exact tradeoff, so that the phase-space density is not changed by this method. However, any "real" ensemble—that is, an ensemble consisting of a finite number of particles or with finite phase-space resolution—is subject to the coarse-grained analysis above. For such a system, the coarse-grained probability density actually drops and the CGE increases during the free-particle evolution. An obvious extreme illustration of this is to consider an initial ensemble of just a few particles: as the faster ones move further away from the center of phase-space, the asymptotic solution has these particles widely spread out in phase-space. During the harmonic oscillator part of the process, all particles travel in circles with the same frequency and there is no change in phase-space density. In Figs. 1–3 we show results from a numerical simulation of the free-particle expansion phase of delta-kick cooling, showing both the reshaping of the distribution as well as the increase in the coarse-grained entropy.

This increase of CGE means that delta-kick cooling is therefore most applicable when starting with significant phase-space density, or equivalently when the parts of the distribution at large q can simply be discarded, in a sense analogous to evaporative cooling, which discards parts of the distribution with high momentum.

We now turn to specific phase-space dynamics illustrations for how there can be a net *decrease* in the CGE using Hamiltonian dynamics.

IV. PHASE-SPACE DYNAMICS AND CGE DECREASE

The previous discussion, including the Rydberg atom results, show that protocols for CGE decrease using Hamiltonian dynamics are both theoretically and experimentally viable and have potentially interestingly applications. It is important to uncover how the details of the Hamiltonians chosen qualitatively and quantitatively affect this coherence enhancement, and how the phase-space structure can be exploited in general. One route to do this is to examine the phase-space trajectory dynamics, that is, the way in which individual trajectories travel on nonlinear orbits, and to see how this controls the global oscillatory behavior of the coarse-grained entropy. Since our interest is in building intuition for the general case, the discussion below is largely qualitative and depends strongly on visualization of trajectory and distribution dynamics in phase-space.

We have studied this with some care for some paradigmatic nonlinear problems, in particular for a one-

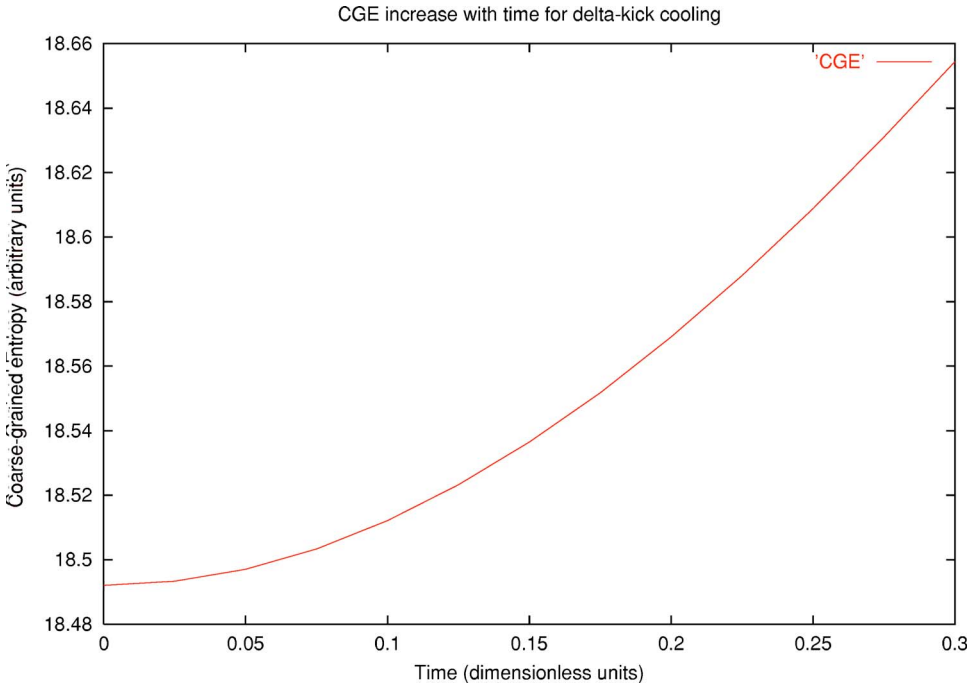

FIG. 1. (Color online) Increase of CGE for delta-kick cooling.

dimensional (1D) sinusoidal potential that is "always on," or applied in short pulses (approximated as delta-function kicks). The always-on system is the pendulum, of course, while the kicked system is known as the Standard Map. The restriction to these 1D problem is useful: First, this enables us to get to the physics without many complications. Second, many interesting experiments involving laser cooling and trapping have actually been done with such 1D systems [26,27]. Our Hamiltonians for the two systems in suitably scaled units are

$$H_{\text{pendulum}} = p^2/2 - \kappa \cos(q), \tag{31}$$

$$H_{\text{Standard Map}} = p^2/2 - \kappa \cos(q) \sum_n \delta(t - nT). \tag{32}$$

The phase-space dynamics for the pendulum for any value of κ can be represented by curves in the q, p plane of two primary kinds. The first are closed, nearly elliptical at lower energy, curves corresponding to oscillatory motion and the second are lines running across the phase-space from left to right for positive p and from right to left for negative p. These latter lines correspond to the pendulum rotating around its pivot point. There is a stable point at $(q=0, p=0)$ consisting of the pendulum being stationary vertical downwards with zero energy [identified with the point $(q=2\pi, 0)$]. An unstable fixed point (a saddle) exists at $(q=\pi, p=0)$ corresponding to the oscillator being vertical upwards. The stable and unstable manifolds associated with this unstable fixed point are the same, and this pair of manifolds forms the separatrix between the two different solutions discussed above for the system. Further, (i) for the oscillating solution, the period of the solution increases with the initial amplitude and/or initial energy and (ii) these trajectories spend a longer proportion of this time in the neighborhood of the turning point as the initial amplitude and/or initial energy is increased; the separatrix has infinite period. These two properties turn out to be very important in focusing. The Standard Map has far more complicated dynamics for general values of κ but for $\kappa \ll 1$, the dynamics are essentially the same as

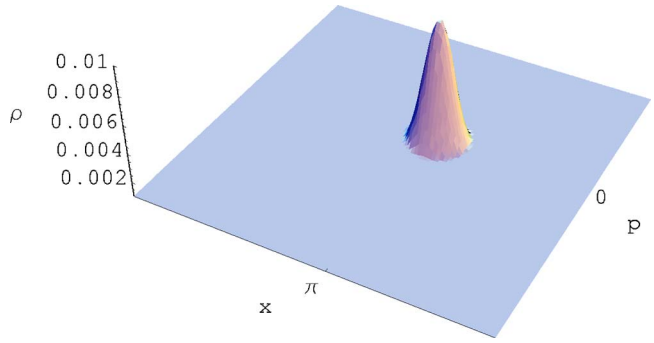

FIG. 2. (Color online) Initial condition for delta-kick cooling.

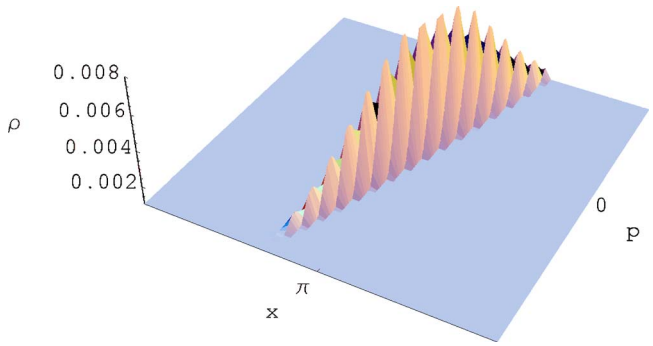

FIG. 3. (Color online) Distribution after free expansion for delta-kick cooling. Notice the change in shape, as well as the decrease in peak height. The jagged character of the distribution comes from the finite number of ensemble members in each bin.

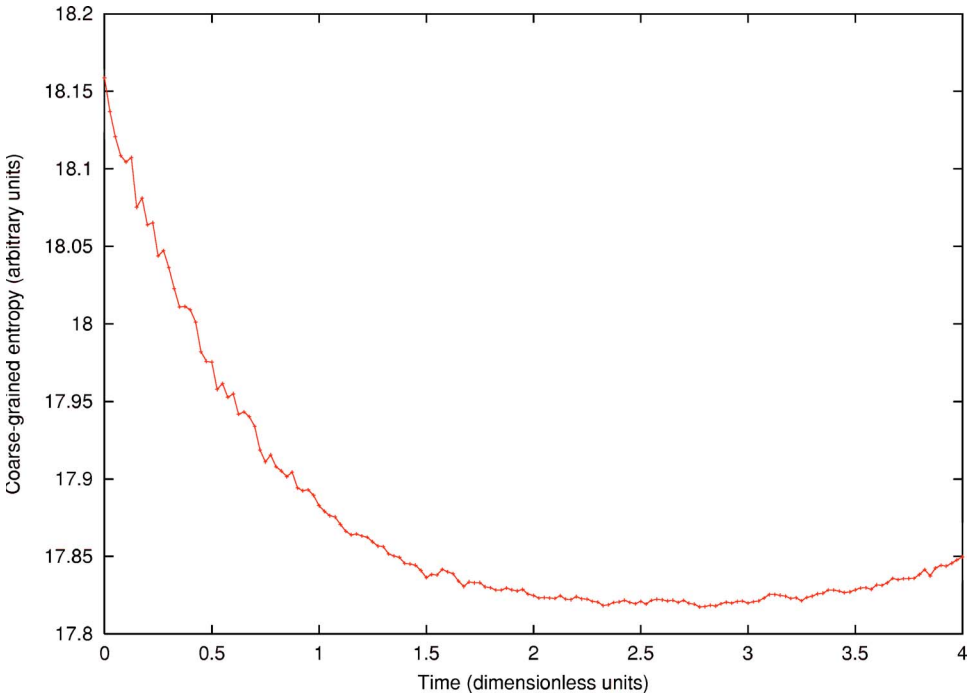

FIG. 4. (Color online) Decrease of CGE followed by its increase for grad-H effect.

for the pendulum, at least for the purposes of our discussion. The point of using the Standard Map is to clarify that kicked systems can be intrinsically Hamiltonian and also to make computation faster in general.

We study probability distributions for this system by specifying many initial conditions and binning them in boxes in the q, p plane, enabling us to construct a CG distribution [7]. $\rho(t)$ is generated by evolving each trajectory and generating the binned distribution again. The coarse-graining neglects higher-order information about ρ; when the nonlinear dynamics moves information up from the finest scales, the information gained about ρ localizes it. We assume that to externally manipulate ρ we can (i) apply the pendulum potential either constantly or in kicks, (ii) and apply nonlinear momentum kicks where the momentum transfer is position dependent. Simpler operations assumed possible are (iii) Free-particle expansion (no potential), (iv) harmonic oscillator evolution as needed. While there is no specific attempt here to consider experimental protocols, we do work with tools that have previously been implemented in atomic systems for example. We have uncovered effects of interest which we now discuss.

A. "Grad-H" effect

The first effect (which we term the "grad-H" effect) arises from the way phase-space trajectories slow down as they approach the turning points of their orbits. As such, turning points or unstable fixed points are natural foci for ensembles, since trajectories spend a lot of time in the neighborhood of such points. The time spent is given by the inverse of the phase-space velocity field at each point. That is, for a general Hamiltonian H, the time spent at a point $(\vec{p},\vec{q})$ is proportional to $|v|^{-1}$ where the phase-space velocity is

$$|v| = |(\dot{\vec{p}},\dot{\vec{q}})| = |\vec{\nabla} H| = \left| \left(\frac{\partial H}{\partial p}\right)^2 + \left(\frac{\partial H}{\partial q}\right)^2 \right|^{1/2}. \tag{33}$$

Consider an ensemble isolated on a microcanonical shell $[H(\vec{p},\vec{q})=E]$ of a nonlinear integrable Hamiltonian (that is, all members of the ensembles have the same energy) and for simplicity in the discussion that follows, consider this to be in a 2D phase-space, so that the microcanonical shell is a closed orbit. It is intuitive that as every trajectory spends greater time in the slower part of the orbit, a cloud of particles started off elsewhere on the orbit will tend to become dense in this part of phase-space.

Formally, consider an element of area $d\omega(0)$ at time $t=0$ on a microcanonical surface. If there is a total of N ensemble members, then the number of these lying within $d\omega$ is

$$dN = N\sigma(0)d\omega(0), \tag{34}$$

where σ is the microcanonical surface probability density for the ensemble (that is, σ is the notation we use for ρ when restricted to a microcanonical surface). If we follow this cloud dN as it evolves, at some later time t_1,

$$dN = N\sigma(t_1)d\omega(t_1) = N\sigma(0)d\omega(0). \tag{35}$$

Now recall that the invariant probability measure for a Hamiltonian on its microcanonical shell corresponding to this area $d\omega$ is the Liouville measure [28]

$$d\mu = \delta[E - H(\vec{p},\vec{q})]d\vec{p}\; d\vec{q} = \frac{d\omega}{|\vec{\nabla} H|}. \tag{36}$$

This being constant, we have the equality

$$\frac{d\omega(0)}{|\vec{\nabla} H|_{t=0}} = \frac{d\omega(t_1)}{|\vec{\nabla} H|_{t=t_1}}. \tag{37}$$

Combining Eqs. (35) and (37), we get that [29],

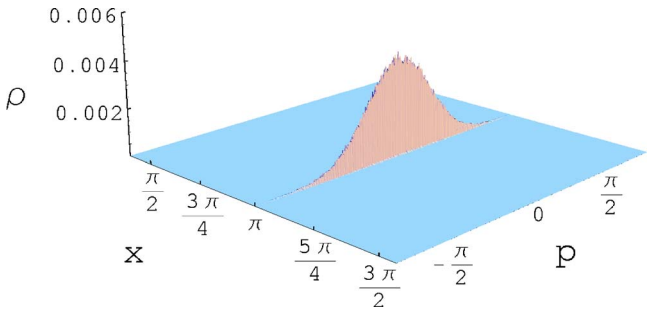

FIG. 5. (Color online) Initial condition for grad-H effect.

$$\sigma(0)|\vec{\nabla}H|_{t=0} = \sigma(t_1)|\vec{\nabla}H|_{t=t_1}. \tag{38}$$

i.e., that the surface probability density alters form as it evolves on the microcanonical trajectory so as to conserve $\sigma|\vec{\nabla}H|$. This can be exploited to transiently enhance the phase-space density $\sigma(t)$ significantly using nonlinear Hamiltonians, where the phase-space velocity varies signficantly in different regions on a microcanonical shell. Specifically, as a result of this grad-H effect, and an ensemble on the microcanonical shell chosen such that it is not yet the invariant distribution, the entropy must oscillate. With the right initialization, the intensity increases periodically as the particles focus near the turning points. In a 2D phase-space, this corresponds to distributing ensemble members at arbitrary phases along a natural orbit and watching them subsequently focus at the turning point.

Note that this effect is then explicitly independent of any coarse-graining for the distribution. The degree of density enhancement depends on the ratio v_i/v_m where v_i represents an average over the absolute value of the initial phase-space velocity and v_m represents the minimum absolute value of the phase-space velocity (at the turning point) along the orbit. For a sufficiently nonlinear orbit and with appropriate choice of initial condition this can be a very large number. In fact, under idealized conditions of the initial condition being a line segment localized along the stable manifold of a saddle point, all the trajectories reach the saddle point in infinite time, giving *infinite density* and zero entropy at this point.

As an application of this method, for example, consider starting with an initial condition that is a flat thin distribution (essentially a line segment) in phase-space. It is now possible

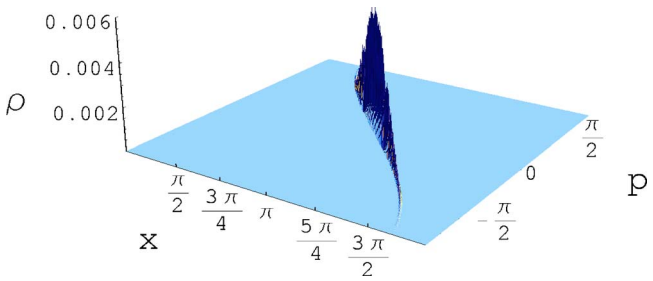

FIG. 6. (Color online) Previous state after being subject to a position-dependent kick that places it on the stable manifold of an unstable fixed point.

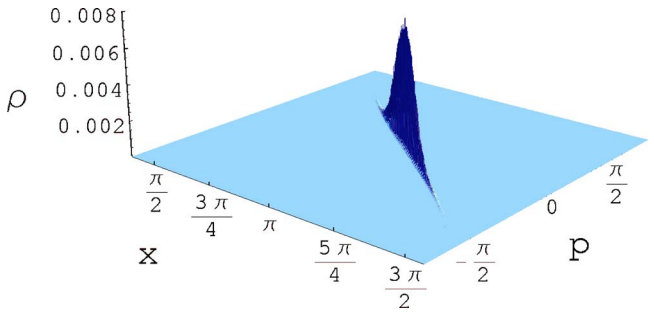

FIG. 7. (Color online) Previous state after evolution in the pendulum potential, at $t=1.0$, when the distribution has contracted along the stable manifold.

to provide a position-dependent momentum kick such that this line is placed on the "fast" part of a phase-space orbit. We would now see that this distribution increases its coherence periodically as it evolved on a given elliptical orbit or ring of orbits. The more this distribution can be made to approximate the fast subset of a natural microcanonical orbit, the more effective the enhancement will be.

We have done numerical experiments to demonstrate these arguments, and some of the results are show in our figures. In Figs. 4–9, we work with a pendulum system, and start with a Gaussian distribution of trajectories which is almost a line segment in phase-space. This state is then given a position-dependent kick so that the central line now approximates a subset of the pendulum separatrix. Given the initial distribution of finite width, it now actually spans a set of different energies near the separatrix as a result. That is, it is distributed across a band of orbits of the nonlinear oscillator. As the distribution evolves, the peak density increases dramatically. We have chosen the central orbit to be the separatrix, which maximizes the increase in density as the entire distribution seems to converge to the unstable fixed point. To understand how this is working even though ρ is not on the microcanonical shell, note that any such distribution is a sum (integral) over distributions, each on a microcanonical shell. Now, since Hamiltonians do not mix contributions from different shells, the argument for each shell applies separately. Since the orbital frequencies are not very different for a narrow enough band of orbits, the different shells will "focus" on approximately the same time scale.

There are some points to note: (i) The time for which the focusing is useful is a function of the width of the Gaussian

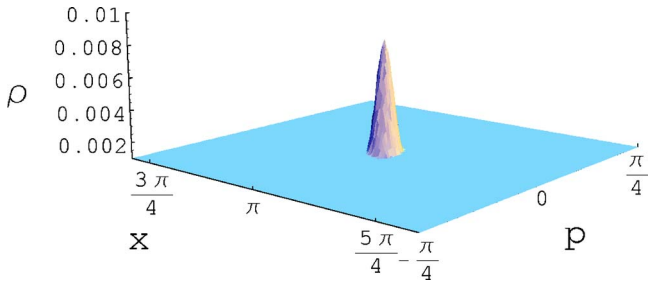

FIG. 8. (Color online) Same as above, at $t=2.5$. Note the change of scale on all three axes. The distribution is now sharply focused at the unstable fixed point.

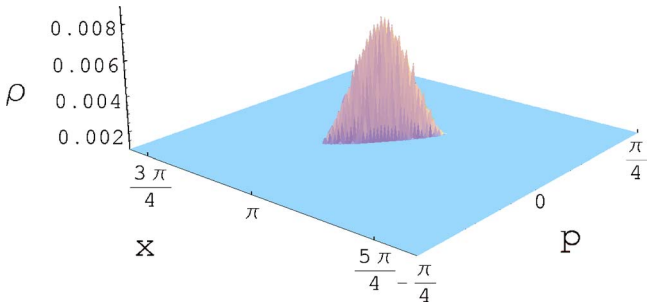

FIG. 9. (Color online) Same as above, at $t=3.25$. After the transient focusing, the distribution begins to stretch along the unstable manifold of the stable fixed point.

that is fed into the process; as the dynamics continue thereafter, the trajectories leave along the unstable direction, resulting in a net decrease of focus. (ii) Such behavior occurs near all unstable fixed points (and a parabolic oscillator potential would work as well, for example). By choosing potentials appropriately, we can manipulate the ensemble as needed. (iii) After focusing, the distribution can be trapped at the stable fixed point of the pendulum by moving the potential and applying a momentum kick as discussed before. (iv) There is an interesting parallel: This effect mirrors in phase-space the techniques [30] that use anomalous diffusion in momentum such that an ensemble spends an increasing amount of time in the $p=0$ region and hence focuses the distribution to lower temperatures.

These results demonstrate that Hamiltonian dynamics appropriately used can lead to phase-space focusing; this method is directly applicable in situations where the experimentally accessible distribution is in the shape of a phase-space cigar, for example.

If the trapping kicks are not applied, in the long run the oscillatory behavior of the distribution on the microcanonical shell dephases. The distribution then relaxes to the invariant distribution for the orbit. In this particular case, the dephasing is expected to happen over many natural time periods of the orbit in question. The dephasing is due to noise effects, computationally corresponding to numerical error, or physically to environmental perturbations or small particle-particle interactions which is why it takes a long time in general. For the pendulum, the invariant distribution has two peaks corresponding to the two turning points for the orbit. This double-peaked distribution is not very useful since it arises at very long times; further, since the distribution fills the entire microcanonical shell, it does not show the same dramatic enhancement of intensity as the short time behavior does, as discussed above. However, even this state can be manipulated further using a different characteristic of nonlinear phase-spaces, as we discuss below.

B. "Inverse dephasing" effect

The grad-H effect, which is most effective on a single microcanonical shell, is not the only useful focusing effect in nonlinear Hamiltonian dynamics. It is possible to increase the coherence even when the distribution is across various different microcanonical shells (as previously shown). Even when given the microcanonical invariant as initial condition (a state that is explicity unaffected by the grad-H effect), it can be further manipulated for enhanced coherence. This is through a second effect we term the "inverse dephasing" or "soundless echo" effect by analogy with the second half of the spin-echo problem.

As previously discussed in the spin-echo problem, in a nonlinear integrable oscillator nearby trajectories travel around in closed orbits with varying frequencies. If a range of these orbits is populated such that the faster orbits are "behind" the slower orbits in their approximately common route, then at some future time the particles will all be closer together in phase-space. This is the basis of what we call the inverse dephasing effect. The difference between our suggestion here for the use of this effect in general Hamiltonian systems and the spin-echo case is that the latter is based on the knowledge that at some time in the past a coherent state existed, so that some form of time-reversal can help access it. In contrast, we will now show how to take some initial conditions and exploit the geometry of phase-space so that the "coherent" or relatively focused state is a future state of the system. The spin-echo effect does not rely on a Maxwell's demon, but a so-called Loschmidt's demon (the macroscopic arrangement of "initial conditions" obtained from an initially coherent superposition via dephasing and reflection to get future coherence). Our suggested protocol is a variation on this demon, one that exploits our knowledge of the nonlinear geometry of phase-space to get future coherence.

This inverse-dephasing or soundless echo effect is exaggerated for nonlinear oscillators by the slow-down (the grad-H effect) near the turning points of the orbits which will all typically be in the same phase-space neighborhood. The spin-echo problem does not show this grad-H effect because the dynamics are those of a linear oscillator with different frequencies for each member of the ensemble.

We have studied this in the Standard Map problem. We take as initial condition the microcanonical invariant distribution on a given orbit with its natural peaks and perturb it by moving it in phase-space relative to the natural orbit. Natural ways of perturbation are (i) rotate the distribution in phase-space using the harmonic oscillator potential—since the initial condition is an invariant for the nonlinear potential which is not a circular shape, this will distribute the ensemble members along various different orbits. (ii) Give the distribution a p kick—that is, give every member of the distribution the same momentum translation (or equivalently, give the distribution a q kick). (iii) Change the kick strength κ slightly, that is, introducing a perturbation $\Delta\kappa$. In principle, we can also combine these perturbations.

In all of these cases, we find that the CGE oscillates in time, and more importantly, shows a very satisfactory decrease during these oscillations. That is, the coherence increases significantly in general. Note that we do not have to start with the microcanonical ensemble, of course. The behavior is essentially the same in all three cases, and we focus in what follows on the particular case of the third perturbation above (changing κ).

For a general argument for time dependence of the CGE, remember that the oscillations come from the average over oscillations as in Eq. (24). Let us take the oscillations of the

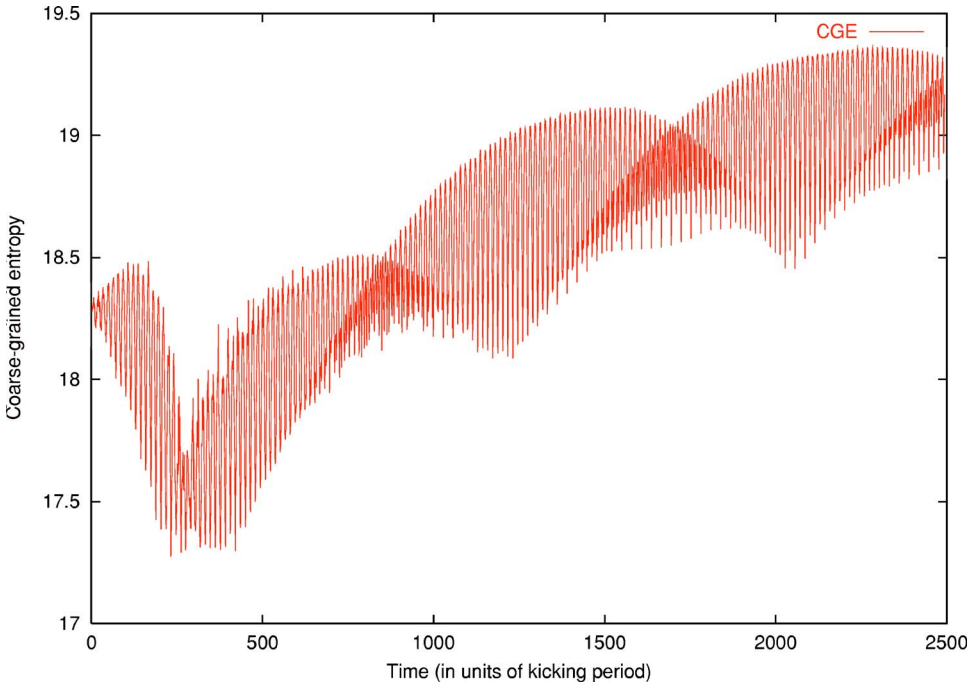

FIG. 10. (Color online) Dephasing oscillations in the CGE for the 'inverse dephasing' effect with $\kappa=0.065$, $\Delta\kappa=0.01$. Notice the first prominent dip in the curve before the oscillating drift higher.

tangent space (or equivalently the distribution gradient) around individual orbits as being approximately sinusoidal $\cos(\omega_i t+\alpha_i)$ where ω_i is the frequency and α_i is the phase of the ith oscillation. The final result, a weighted average over many different oscillations, can be approximated to first order to be of the form

$$\sum_i A_i \cos(w_i t + \alpha_i), \tag{39}$$

where A_i is the weighting factor (in general, of course, the time dependence is oscillatory but not sinusoidal). This sum is therefore finally of the form

$$\chi^2 \simeq \cos(\bar{\omega}t)\cos(\Delta\omega t), \tag{40}$$

where $\bar{\omega}$ is the frequency averaged over the different orbits and $\Delta\omega$ is the spread in frequencies. Since $\Delta\omega$ is typically much smaller than $\bar{\omega}$, this leads to the prediction of "rapid oscillations" at $\bar{\omega}$ with an envelope function that is oscillating at $\Delta\omega$. Also, as a result of the combination of the two oscillations, the peak coherence obtains through the $\cos(\Delta\omega t)$ term and hence should be on the time scale $\tau \approx 1/(\Delta\omega)$.

We have seen these kinds of behavior clearly in the numerical experiments. The results show that the CGE as a function of time exhibits dephasing oscillations that are very analogous to beating. The behavior is completely general for a very large range of parameters, differing only in specific details of size and time scales. We show typical oscillations for an initial $\kappa=0.065$ and a $\Delta\kappa=0.01$ in Fig. 10 as well as snapshots of the distribution in Figs. 11–15 as a function of time, showing the clear change in phase-space coherence and corresponding increase in peak height. The snapshots of the distributions when followed over long times show a similarity to the "revival" part of the quantum phenomenon of "wave-packet revival" [31]: The difference, as with the spin-echo problem, is that in this result we are showing the creation of a coherence, and not the recovery of a coherence.

Note that as the perturbation strength $\Delta\kappa\to 0$, the number of different orbits populated also $\to 0$ and hence $\Delta\omega\to 0$, which implies that the time to the maximum $\tau\to\infty$, as can be seen in Fig. 16. A similar effect is seen in the the scale of the

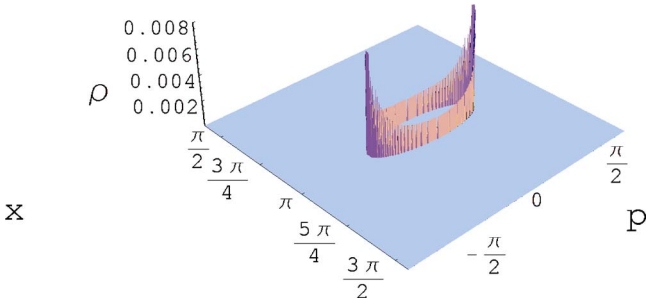

FIG. 11. (Color online) Initial condition for the inverse dephasing effect, the microcanonical distribution for an energy shell, with a double-peak corresponding to the two unstable fixed points.

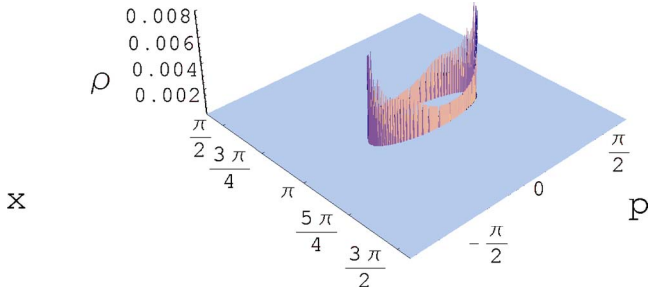

FIG. 12. (Color online) At $T=190$ kicks, a transient stage verifying that because of the various frequencies participating, the distribution can actually be less focused on the way to greater focusing.

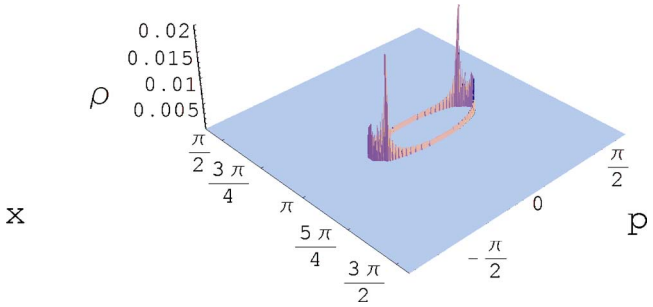

FIG. 13. (Color online) At $T=320$ kicks, the peaks are clearly visible already, and can be seen to be not at the turning points.

increase of coherence, shown in Fig. 17. Again, both the grad-H and inverse dephasing effects contribute to this. As $\Delta\kappa \to 0$ and $\Delta\omega \to 0$, the coincidence of the different periods becomes better, leading to an enhanced increase of coherence for decreased perturbation in general. However, while $\Delta\omega$ monotonically decreases with $\Delta\kappa$, the ensemble distributes itself in complicated ways on the different orbits as a result of the perturbation. That is to say, the coefficients A_i from Eq. (39) are a complicated function. The details depend on the detailed shape of the relevant orbits, and the form of the initial distribution. The maximum decrease in entropy $\Sigma = \Delta S_{\max}$ is therefore not a monotonic function of $\Delta\kappa$, even though there is a general trend to better results at small perturbations.

Loosely speaking therefore, the best results are obtained for the smallest perturbation (although this has to be followed by the longest wait). This rule of thumb is constrained by the impact of the ultimate resolution limit (the finite size of the coarse-graining cells) as well as the dynamical coarse-graining (numerical round-off error). That is, there is a saturation effect for the smallest perturbations.

To summarize the connection between phase-space structure and CGE dynamics therefore:

(i) If mapped onto an appropriate subset of a microcanonical distribution for a nonlinear Hamiltonian, the grad-H effect works to enhance the density. As appropriate, the more extreme the change in phase-space velocity on an orbit, the larger this effect.

(ii) An initial condition that resembles the microcanonical distribution of a specific nonlinear Hamiltonian can be perturbed onto the neighboring energy shells to exploit the inverse dephasing effect. The smaller the perturbation, the

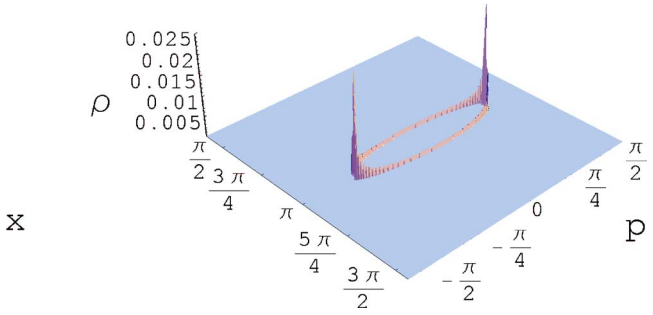

FIG. 14. (Color online) At $T=333$ kicks, at or near a maximum in the transient focusing from inverse dephasing.

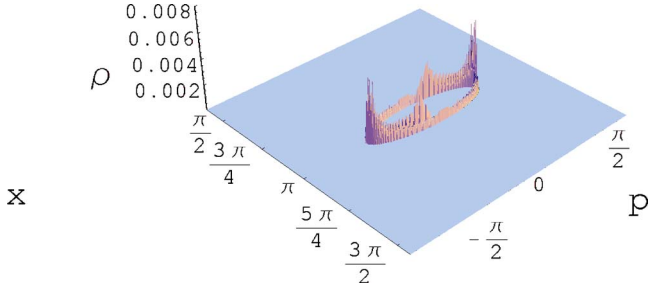

FIG. 15. (Color online) The distribution from Fig. 14 seen much later (at $T=2400$ kicks) showing that the dephasing has lead to an incoherent distribution spread across the accessible orbits.

better the coherence enhancement, although the longer it takes. There is a limit to this, arising from noise effects.

V. DISCUSSION AND CONCLUSIONS

We now consider some applications resulting from the coarse-grained analysis.

A. Violating Liouville's theorem in infinite time

One point of theoretical interest is that by combining ideas from delta-kick cooling and the grad-H effect on an unstable manifold, Liouville's theorem can be seemingly violated. To see this, consider starting with a standard two-dimensional distribution, a Gaussian in phase-space space, say. If we allow this distribution to evolve under the free-particle Hamiltonian, an infinitely thin Gaussian will be obtained in infinite time. This transforms a two-dimensional distribution with finite thickness in all directions to one that is infinitely thin in one direction and infinitely long in the other direction. Now consider this one-dimensional distribution being rotated onto the stable manifold of an unstable fixed point, that is, applying a harmonic oscillator potential for the appropriate amount of time, followed by a static inverted quadratic potential $V(q)=-q^2/2$. In that case, after infinite time the entire distribution will accumulate at the origin, thus being reduced to a zero-dimensional distribution. We therefore seem to have violated Liouville's theorem, changing the phase-space volume from something finite to zero.

Liouville's theorem remains valid for practical purposes, however, since this result requires two procedures each taking infinite time that must be followed in sequence. If either or both processes is followed for finite times, the free-particle evolution results in a finite-width Gaussian that can then re-focus using the stable manifold, but will yield a two-dimensional distribution that is only as focused as the initial condition itself.

B. Improved cooling

A simple practical application of these ideas is to improve the delta-kick cooling protocol. Remember that delta-kick cooling consists of two phases, the first being a free-particle expansion phase which converts an arbitrary Gaussian in the $q-p$ phase-space into a long thin version (a cigar shape)

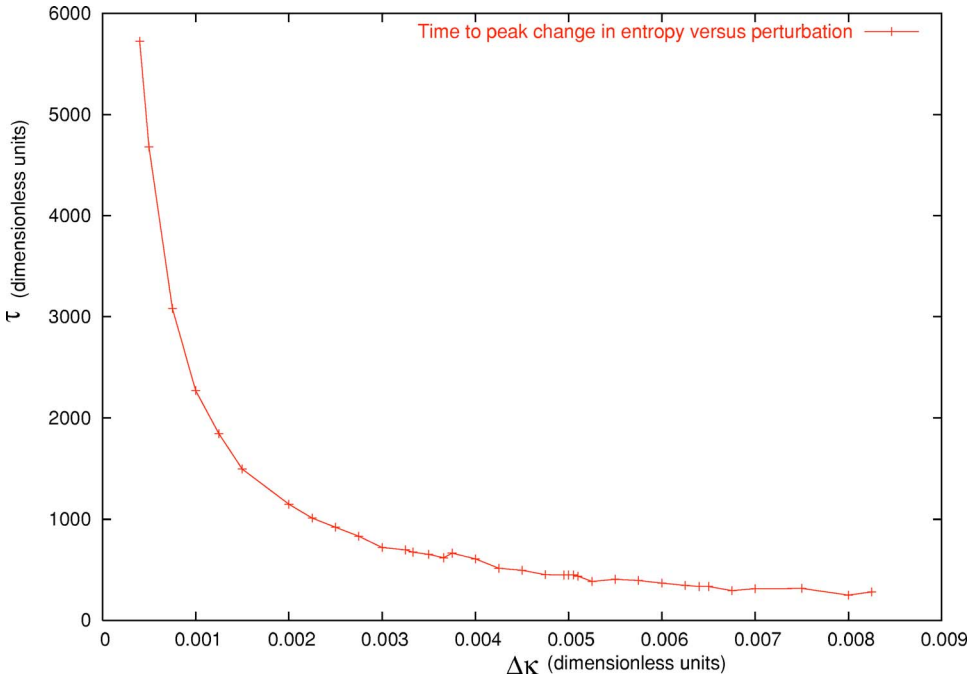

FIG. 16. (Color online) Time to peak in CGE versus perturbation for inverse dephasing effect.

oriented at an angle to the q axis. This angle depends upon the time for which the free-particle evolution takes place, and the second phase of the protocol is to rotate the distribution onto the q axis by applying the harmonic oscillator potential for the appropriate length of time, which depends on the length of time used for free-particle evolution.

We suggest replacing the free-particle evolution with evolution in the inverted quadratic potential. After a rapid transient initial evolution inwards along the stable manifold of the unstable fixed point of that potential, the Gaussian then evolves to stretch along the unstable manifold of the potential, rapidly becoming infinitely thin and long. This means that the time to a given temperature or momentum width can be reduced dramatically from the free-particle case, and can be made as short as needed, by increasing the curvature of the potential. Second, the time for which the static harmonic oscillator potential is applied is now a fixed quantity, since it depends only on the curvature of the potential. This reduces the need to measure or control the time of free-particle evolution to high accuracy.

Another point is that instead of changing various potentials for specified periods of time as above, one could instead use an "always-on" potential with q^n where n is a high even number $(n=10)$ say. This yields dynamics essentially the same as delta-kick cooling—a reasonably well-localized distribution experiences q^n for $q \ll 1$ and is initially essentially

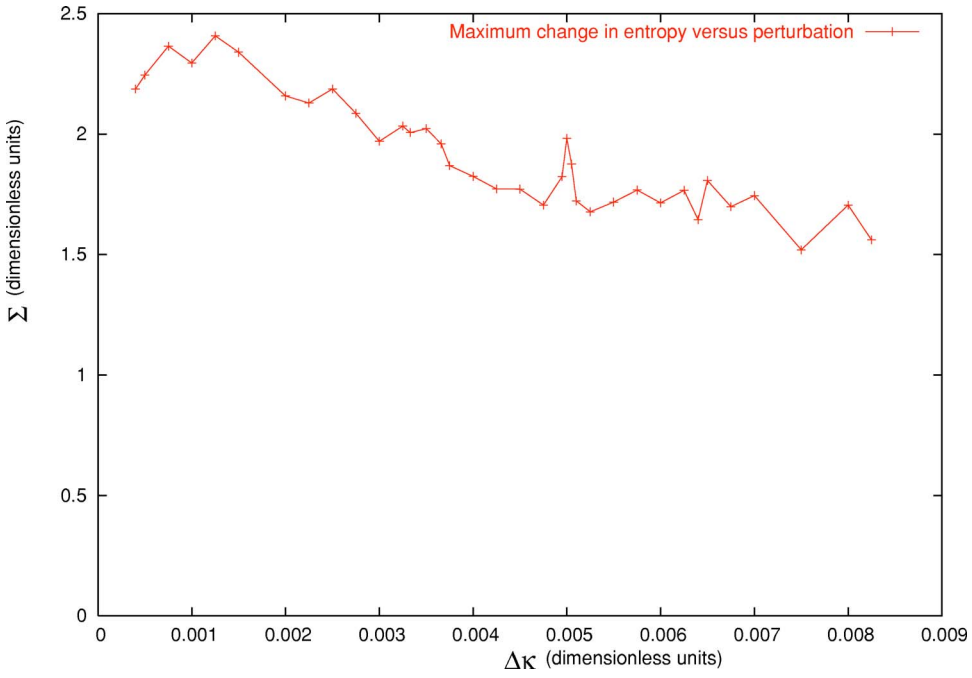

FIG. 17. (Color online) Maximum in change in CGE versus perturbation for inverse dephasing effect.

free. It therefore stretches into a cigar shape as all the particles travel parallel to the q axis with speeds that depend on their distance from the axis. As the faster particles get to the $q \gg 1$ regime, they encounter a sharp potential wall, and change direction abruptly, which results in the distribution turning sharply in phase-space to orient along the q axis. The specific dynamics in this case yields a small curlicue at the ends of the distribution coming from the nonlinear turning, although the effect of this curlicue can easily be made negligible.

As argued above, however, delta-kick cooling is in general accompanied by an increase in CGE and hence a decrease in effective phase-space density. We have shown above that it is possible to decrease CGE and increase effective phase-space density for certain situations. In general, there is no formal reason why it should not be possible to find a set of potentials and corresponding protocols such that the distribution cannot have its CGE decreased. The distributions that are most easily and maximally improvable look like simple tendrils in phase-space (that is, long and thin distributions which are not wrapped in complicated ways since that would make the potentials and protocols for improvement equally complicated).

C. Gaussians moved between potentials of different curvature

Clearly, we cannot decrease the CGE when the initial distribution has the minimal possible CGE for that particular distribution. To understand the idea of the minimal CGE state, remember that the CGE is extremely well-described by the (hyper)surface area for the distribution. For a distribution of given phase-space (hyper)volume, this means that there is a shape that minimizes its CGE. This is the shape of maximum symmetry, the sphere: the minimum perimeter for a given area in two dimensions is for a circle.

A thermal Gaussian can be represented, in suitably scaled units, as a circle in phase-space. Therefore the above seems to imply that it is not possible to decrease the CGE for a distribution that is initially a circle and hence that a thermal cloud cannot be further focused. However, this turns out to be true only if we continue to work with the same harmonic oscillator throughout. If we start with a "circular" Gaussian distribution generated in a given harmonic oscillator, it can be manipulated to decrease the CGE if the final target state is for an oscillator with a different frequency.

The argument goes as follows: Working in two phase-space dimensions for simplicity, consider any harmonic oscillator and a thermal Gaussian state as initial condition. Since phase-space has no well-defined metric, we can now rescale our q, p units such that this distribution is represented as a circle. As such, the distribution would then have "minimum perimeter" for its area. But this cannot be enough to define the lowest entropy state, since a Gaussian of different shape can also be scaled to be circular with a different system of phase-space units. We therefore need additional criteria for a unit of minimal entropy. One possibility for such an object is the minimum-uncertainty Gaussian for a given harmonic oscillator (also known as a coherent state in quantum mechanics, equivalent to a possibly displaced ground state). This object is symmetrical in q, p and can hence be used to define a circle in phase-space. Interestingly, this state also does not change shape under the action of the harmonic oscillator Hamiltonian. Any other Gaussian in phase-space corresponds to a squeezed state: These "breathe" as a function of time where they continue to have the same area, and hence the same uncertainty, but change shape. This makes using a minimum-uncertainty state sensibly self-consistent. Such a minimum-uncertainty Gaussian is therefore the correct candidate to define a circle for a given harmonic oscillator. All Gaussians with the same shape (aspect ratio in p, q, as well as lack of $p-q$ correlation) as a minimum-uncertainty state then correspond to circles in phase-space; this allows them to be of arbitrary size.

However, this means that every harmonic oscillator has a different definition for a circle or for a minimum entropy state. A state that corresponds to a circle for a given harmonic oscillator looks like an ellipse for an oscillator with a different frequency. This suggests that we can change the CGE for an initial condition that is created as a thermal Gaussian within a given harmonic oscillator, say, by changing the curvature of the external potential—that is, by placing it in a different harmonic oscillator. It might seem paradoxical that the minimum entropy state is not absolute. That is, since every harmonic oscillator has its own definition for a minimum entropy state, we cannot tell if a state is at minimum area without specifying which harmonic oscillator we are considering. Some consideration shows that this is not a contradiction, however. We have to specify our resolution (quantum mechanically, the target wave function with which we are going to overlap our distribution) before deciding how something can be observed. This is precisely the same meaning as these ideas about Gaussians in different oscillators.

As a practical result, it means that we can use CGE-analysis-based protocols as above to both decrease temperature (as in delta-kicked cooling) and increase phase-space density for Gaussian initial conditions. In particular, if the initial condition can be treated as a "vertical ellipse" then we can both increase the density and reduce its momentum spread in making it a circle by working with a target state in a harmonic oscillator of greater frequency.

D. "Spin-echo effect" in atomic ensembles

The somewhat counterintuitive "refocusing" of the signal in the spin-echo effect has generated debate in the literature on the nature of reversibility in physical systems. This partly stems from the lack of many other experimentally implemented examples of this kind of behavior [17]. The critical issue in this effect is that the refocusing in microscopic dynamics is effected by macroscopic changes in external potentials. Our discussion above shows that the spin-echo system is not intrinsically distinct from other Hamiltonian systems. As such, in the interest of clarifying this discussion of reversibility and coarse-grained entropy, we point out that it is straightfoward to show the equivalent of the spin-echo effect in atomic ensembles.

Possible protocols, among the many that can be imagined, are as follows:

(i) Start with any finite ensemble, possibly distributed as a Gaussian, localized at $q=0$, $p=0$ in phase-space.

(ii) Allow for free-particle evolution (equivalent to the free precession part of the spin-echo dynamics) for time $t=T_1$. As the distribution evolves, the central density and experimental signals associated with the density degrade. Consider a straight line drawn to a particle originally on the p axis; as a result of the evolution, this is now oriented with respect to the q axis at some angle θ_1 which depends upon T_1. Alternatively, as suggested above for the improved delta-kick cooling situation, allow the distribution to evolve on the inverted quadratic potential; in appropriately scaled units, in this case, the equivalent angle is $\theta_1'=\pi/4$.

(iii) At this stage, equivalent to the spin-flip or magnetic-field flip part of the spin-echo problem, apply a harmonic oscillator potential for a time $t=T_2$. If (ii) was free-particle evolution, then this has to be chosen so that the distribution is rotated clockwise through an angle $\theta_2=2\theta_1+\pi$. If (ii) was on the inverted quadratic potential, the clockwise rotation should be through $3\pi/2$.

(iv) This is the stage equivalent to the refocusing part of the spin-echo problem. Depending on the choice in (ii) again, we now allow for free-particle evolution again or for evolution on the inverted quadratic potential, respectively. As in the spin-echo problem, the ensemble will transiently refocus as a result of the evolution [32]. The experimental signal will increase in strength as a result of the refocusing.

This protocol can be easily implemented in the laboratory, and should help clarify that the spin-echo effect is not intrinsically different. Further, decays in the experimental signal as a function of the time spent in (ii) would be an interestingly simple monitor of environmental and many-body effects.

E. Possible quantum effects

How is all this affected by quantum effects? The broader ideas of changing coarse-grained entropy and corresponding changes in the physical distribution clearly hold even when the evolution is quantum mechanical.

Further, the delta-kick cooling protocol presented here, using (at most) quadratic potentials, including the modifications that we have introduced above, are completely unaffected when we consider quantum corrections. This is because classical and quantum distributions evolve identically in quadratic potentials.

The grad-H effect is the most challenging to translate into quantum mechanics because of the occasionally remarkable differences in quantum and classical behavior in the neighborhood of unstable fixed points in phase-space. However, this effect can be implemented using the inverted quadratic potential as well. In that case again, the quantum dynamics follow the classical ones, and the predictions carry over.

Finally, the particular issue of the "inverse dephasing" effect: In the classical case, there is a continuous spectrum, that is, the classical orbits have a continuum of frequencies, whereas in the quantum case, there are only a discrete set of frequencies involved. This means that in the absence of external noise, the CGE oscillations do not dephase quantum mechanically without external effects. This predicition has actually been verified in the Rydberg atom case using quantum calculations [3].

F. Conclusion

Coherence increase is a completely general effect in the behavior of ensembles in nonlinear Hamiltonian systems with potentially interesting applications. By considering simple nonlinear systems, we have made progress in understanding general principles of how this phenomenon works, in particular clarifying the role of classical turning points, and the impact of small perturbations on microcanonical ensembles.

There are some interesting ideas about the efficacy of Hamiltonian dynamics in changing phase-space density or the entropy resulting from this analysis. The "standard" interpretation is that these things (or the degree or localization or focusing) of a classical or quantum density cannot be altered, and in particular cannot be increased, by purely Hamiltonian methods. The phenomenon of coherence enhancement shows that this intuition does not apply in general. In this paper we have discussed at least two ways in which this is altered which can be understood in terms of phase-space trajectories for nonlinear Hamiltonians.

The first is that when thinking of an ensemble on a microcanonical shell, for example, the correct quantity that is conserved is the Liouville measure, which includes a weighting by the local phase-space velocity. As such, evolution on microcanonical shells for nonlinear Hamiltonians with a wide variation in local phase-space velocities can be used to significantly enhance densities. The second is that the shape of distributions matters a great deal in determining the effective phase-space density or localization of the distribution. This leads to the utility of effects such as the inverse dephasing phenomenon, where a phase-space distribution can be placed with the faster orbits behind the slower orbits such that the distribution focuses at a future time.

In conclusion, these analyses confirm that the phenomenon of coherence enhancement in the dynamics of probabilities evolving under nonlinear Hamiltonians is not a trivial or formal effect. The nonlinearity has a nonintuitive impact on their behavior and can be chosen to alter the coherence or degree of localization of the ensemble. We anticipate various interesting applications of this phenomenon; it would be particularly challenging to understand the particular choice of Hamiltonians for given application and initial condition that will allow the effect to be maximized, and this constitutes ongoing research.

ACKNOWLEDGMENTS

This work was supported by a Cottrell College Science Award from the Research Corporation. Students were also supported by the Howard Hughes Fund administered through Carleton College. Part of the research was done during visits by A.K.P. to the Kavli Institute for Theoretical Physics (UC

Santa Barbara) as part of the Quantum Gases program, as well as to the Center for Advanced Studies at the University of New Mexico—it is a pleasure to acknowledge their hospitality. A.K.P. is also grateful to Vladan Vuletic and Nelson Christensen for extended discussions about cooling, and Bruce Thomas for discussions on the entire paper. Comments by Susan Coppersmith on preliminary results helped make the connection to spin-echo dynamics.

[1] See, for example, J. R. Dorfman, *An Introduction to Chaos in Nonequilibrium Statistical Mechanics* (Cambridge University Press, New York, 1999).
[2] C. L. Stokely *et al.*, Phys. Rev. A **65**, 021405(R) (2002).
[3] D. G. Arbó *et al.*, Phys. Rev. A **67**, 063401 (2003).
[4] See, for example, M. Vengalattore *et al.*, Phys. Rev. Lett. **92**, 183001 (2004).
[5] A. K. Pattanayak, Physica D **148**, 1 (2001).
[6] See, for example, the discussion of Renyi information in C. Beck and F. Schlogl, *Thermodynamics of Chaotic Systems* (Cambridge University Press, New York, 1993).
[7] R. Tolman, *The Principles of Statistical Mechanics* (Dover, New York, 1979).
[8] K Takahashi, Prog. Theor. Phys. Suppl. **98**, 109 (1989).
[9] A. K. Pattanayak and P. Brumer, Phys. Rev. E **56**, 5174 (1997); Yuan Gu, Phys. Lett. A **149**, 95 (1990).
[10] W. Liu and G. Haller, Physica D **188**, 1 (2004).
[11] D. Rothstein, E. Henry, and J. P. Gollub, Nature (London) **401**, 770 (1999), and references therein.
[12] R. T. Pierrehumbert, Chaos, Solitons Fractals **4** 1091 (1994).
[13] A. K. Pattanayak and P. Brumer, Phys. Rev. Lett. **79**, 4131 (1997); A. K. Pattanayak, *ibid.* **83**, 4526 (1999).
[14] A. K. Pattanayak, B. Sundaram, and B. D. Greenbaum, Phys. Rev. Lett. **90**, 014103 (2003).
[15] P. Gaspard, *Chaos, Scattering and Statistical Mechanics* (Cambridge University Press, New York, 1998).
[16] A. K. Pattanayak and P. Brumer, Phys. Rev. Lett. **79** 4131 (1997).
[17] For a recent careful consideration of this effect, see D. A. Lavis, Found. Phys. **34**, 669 (2004); eprint cond-mat/0311527.
[18] See the discussion of the Loschmidt echo in the literature, for example, P. R. Levstein, G. Usaj, and H. M. Pastawski, J. Chem. Phys. **108**, 2718 (1998).
[19] J. Ford, G. Mantica, and G. H. Ristow, Physica D **50**, 493 (1991).
[20] V. Latora and M. Baranger, Phys. Rev. Lett. **82**, 520 (1999).
[21] R. S. Dumont and P. Brumer, J. Chem. Phys. **87**, 6437 (1987).
[22] D. G. Arbó, C. O. Reinhold, and J. Bürgdorfer, Phys. Rev. A **69**, 023409 (2004).
[23] W. Ketterle and D. E. Pritchard, Phys. Rev. A **46**, 4051 (1992).
[24] See the detailed discussion in W. Greiner, *Classical Mechanics: Systems of Particles and Hamiltonian Dynamics* (Springer, New York, 2003); and also M. G. Raizen *et al.*, Phys. Rev. A **58**, 4757 (1998).
[25] S. Chu *et al.*, Opt. Lett. **11**, 73 (1986).
[26] H. Ammann and N. Christensen, Phys. Rev. Lett. **78**, 2088 (1997); S. H. Myrskog *et al.*, Phys. Rev. A **61**, 053412 (2000).
[27] See, for example, F. L. Moore *et al.*, Phys. Rev. Lett. **75**, 4598 (1995); and D. A. Steck *et al.*, *ibid.* **88**, 120406 (2002).
[28] Although the last equality is rarely emphasized in textbooks, see, for example, O. de Almeida, *Hamiltonian Systems: Chaos and Quantization* (Cambridge University Press, New York, 1988), p. 60.
[29] B. I. Henry and J. Grindlay, Can. J. Phys. **75**, 539 (1997).
[30] F. Bardou *et al.*, *Lévy Statistics and Laser Cooling* (Cambridge University Press, New York, 2002).
[31] See M. A. Doncheski and R. W. Robinett, Phys. Rep. **392**, 1 (2004), for example.
[32] For distributions initially symmetric in q the signal can be refocused with rotations that can ignore a π reflection during stage (iii) above. For the free-particle case, this means that the rotation would be $2\theta_1$, and for the inverted quadratic potential, the rotation would be $\pi/2$.