Entropy evolution as a guide for replacing the Liouville equation

Brad Lee Holian

Los Alamos National Laboratory, University of California, Los Alamos, New Mexico 87545

A replacement is proposed for the usual Liouville equation for the evolution of the distribution function of a statistical ensemble of classical, homogeneously thermostatted systems. The correct description of the entropy evolution is an essential criterian for such a replacement to satisfy, since we show from numerical realizations of a physically realistic model system that the Liouville equation is *not* satisfactory in general. On the other hand, both the Liouville equation and the proposed replacement correctly describe the evolution of ordinary mechanical observables, such as the energy. The particular form of the proposed replacement is shown to describe correctly a variety of experimental conditions.

I. INTRODUCTION

Classical statistical mechanics is founded on the notion that the calculation of the average value of an observable mechanical property, such as energy or pressure, is somehow made easier by considering a large collection, or statistical ensemble of experiments, with identical boundary conditions but differing in initial conditions. The dynamical behavior of this ensemble of trajectories in phase space is embodied in a time-dependent continuous distribution function, much like the mass density in fluid mechanics. The equation of motion in phase space for the distribution function, traditionally assumed to be the continuity, or Liouville equation, provides a connection between the microscopic equations of motion of individual atoms in a given member of the ensemble to the macroscopic equations of motion for the ensemble average of the observable.

A fundamental criterion for the equation describing the time evolution of the distribution function is that the time derivative of the ensemble average of a typical observable be equal to the ensemble average of its time derivative. This important property of the distribution function and *its* equation of motion is far from simply being a tautology. It allows us to transform from the picture based on the observable—the so-called Heisenberg or Lagrangian view—to the picture based on the distribution function—the so-called Schrödinger or Eulerian view. This transformation is the starting point for classical response theory both linear and nonlinear.[1]

In the Heisenberg picture, we imagine following systems along their trajectories in phase space. Each system is like a Lagrangian "mass" element whose weight is fixed at time $t = 0$ by the distribution function times the differential phase-space volume element it occupies. Along the trajectory, the observable changes with time. The local value of the distribution and the size and shape of the volume element change with time along the trajectory in such a way that the product, or mass of each trajectory is constant in time. The ensemble average of the observable at any instant of time is then its sum over all trajectories, each weighted by its constant trajectory mass.

In the Schrödinger picture, we imagine a fixed mesh of cells in phase space. Positioning ourselves at one such cell, we observe trajectories flowing through the cell like noninteracting "particles" in a multidimensional "fluid." For that cell, the observable is a quantity depending only on the value of the phase-space variables (particle coordinates and momenta) at the center of the cell. The flow of trajectories is reflected in the time-varying distribution function or local probability density, which, when multiplied by the differential volume element and the local value of the observable, and summed over all cells, gives the ensemble average.

The equivalence of Heisenberg and Schrödinger pictures for the ensemble average would seem to be intuitively obvious. In fact, the equivalence is the result of the commutativity of the two operations of total time derivative and ensemble average. Consequently, the Heisenberg-Schrödinger equivalence for ensemble averages is actually a nontrivial property of the equation describing the time evolution of the distribution function. In Ref. 2 we showed that the Liouville equation satisfies this fundamental criterion numerically for a real ensemble of systems. The Liouville equation cannot easily be tested by itself, but properties such as moments of the phase-space distribution can be measured and compared with predictions based on the Liouville equation.

The system studied in Ref. 2 and also in this paper is the one-dimensional thermostatted harmonic oscillator. In addition to the usual coordinate q and momentum p, the phase space $\Gamma = (q, p, \zeta)$ of the Nosé-Hoover oscillator[3] includes a dimensionless friction coefficient ζ which couples the oscillator to a thermal reservoir by a nonlinear, deterministic feedback mechanism. There are two parameters associated with this thermostat: ν, the characteristic thermostatting rate, and T, the temperature, which is proportional to the long-time average of the kinetic energy. This is the simplest example of deterministically thermostatted systems, which have proved to be a powerful theoretical and computational tool.[4,5] (In this paper, we do not restrict ourselves unnecessarily to equations of motion derivable from a Hamiltonian.)

This system provides a good means of testing the com-

mutativity property since a large ensemble can be followed for long times, both at and away from equilibrium. For example, as we show in Appendix A, the Nosé-Hoover thermostatted equations of motion preserve the canonical ensemble distribution; that is, the moments of the Nosé-Hoover oscillator distribution should remain Gaussian in time if the initial conditions are selected from the equilibrium distribution. We found that they do so, even though a large proportion of the trajectories follow regular, nonergodic orbits. Thus, the ability of the Liouville equation to predict the macroscopic equation of motion of an observable does not depend at all on the ergodicity of the system.

A more stringent test of the Liouville equation involves the process of thermal equilibration from initial distributions that are characterized in width by a temperature T_0 different from the thermostat temperature T. If $T_0 < T$, the thermostat heats up the ensemble—the so-called "Big Bang" experiment because of the expansion of the initial distribution. If $T > T_0$, the thermostat cools down the ensemble, causing the distribution to shrink, hence the "Big Shrink." In both cases, the ensemble distribution approaches a second-law attractor, that is, the small but probable region of phase space of the size characterized by a Gaussian of temperature T.

The energetics of the two equilibration experiments are described well by the Liouville equation, as would be supposed from our discussion of ordinary mechanical observables. However, a curious behavior was noted for the entropy, as defined in Sec. II by Eq. (11). If the Liouville equation were correct, the entropy so defined for either equilibration process would never change from the initial value. From our description of the second-law attractor, it is clear that qualitatively the entropy in both equilibration processes rises from a low initial value toward a maximum at equilibrium.

In this paper, we present results for the numerical realization of the entropy in these equilibration experiments. Because the equilibrium distribution for the Nosé-Hoover oscillator ensemble is a simple Gaussian in each of the three Γ-space variables, it is straightforward to construct Schrödinger (Eulerian) boxes in phase space, utilizing the normal error function (see Appendix B), such that on the average there is one ensemble member per cell at equilibrium. This discrete evaluation of the entropy evolution emphasizes that entropy is inherently a Schrödinger property of the ensemble. Put another way, the ensemble average of an ordinary observable is a linear functional of the distribution function; hence, it is possible to show a Heisenberg-Schrödinger equivalence for such observables—we may either follow the flow of trajectories through fixed cells in phase space or follow a distorting mass element along a trajectory. Because entropy is a *nonlinear* functional of the distribution, no such Heisenberg-Schrödinger equivalence is possible—also, entropy is *nonlocal*: there is no entropy function along a trajectory like energy or pressure.

Guided by these numerical results and using the entropy evolution as a second criterion (as well as the commutativity property for ordinary observables), we propose a replacement for the Liouville equation for the dynamics of the distribution function. The additional term is diffusive normal to the usual (q,p) subspace, i.e., in the direction of the friction coefficient ζ. The phase-space mixing is driven by the difference between the nonequilibrium and equilibrium distribution functions at a rate that is similar to the Lyapunov exponent. The proposed replacement of the Liouville equation is then tested against a variety of experimental conditions and found to be superior because of the entropy evolution criterion.

II. MATHEMATICAL CONSIDERATIONS

The first criterion that must be satisfied by a viable candidate to replace the Liouville equation is that the operations of time derivative d/dt and ensemble averaging $\langle \cdots \rangle$ commute. For an observable B which does not explicitly depend on time, but only on the phase Γ, the ensemble average in the Schrödinger (Eulerian) picture is defined by

$$\langle B(t) \rangle = \int d\Gamma f(\Gamma,t) B(\Gamma) , \tag{1}$$

where $f(\Gamma,t)$ is the distribution function. The macroscopic equation of motion, or time derivative of the ensemble average is

$$\frac{d}{dt} \langle B(t) \rangle = \frac{d}{dt} \int d\Gamma f(\Gamma,t) B(\Gamma)$$
$$= \int d\Gamma B(\Gamma) \frac{\partial}{\partial t} f(\Gamma,t) . \tag{2}$$

From the equations of motion $\dot{\Gamma}(\Gamma,t)$, the total time derivative of B at the phase point Γ is

$$\dot{B}(\Gamma,t) = \frac{d}{dt} B(\Gamma) = \left[\frac{\partial}{\partial t} + \dot{\Gamma}(\Gamma,t) \cdot \frac{\partial}{\partial \Gamma} \right] B(\Gamma)$$
$$= \dot{\Gamma} \cdot \frac{\partial B}{\partial \Gamma} , \tag{3}$$

since $\partial B/\partial t = 0$. Hence, the ensemble average of $\dot{B}$ is

$$\langle \dot{B}(t) \rangle = \int d\Gamma f(\Gamma,t) \dot{B}(\Gamma,t) = \int d\Gamma f \dot{\Gamma} \cdot \frac{\partial B}{\partial \Gamma} . \tag{4}$$

If d/dt and $\langle \cdots \rangle$ commute, then Eqs. (2) and (4) are equal.

The Liouville equation is the continuity equation for the distribution function:

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \Gamma} \cdot (f \dot{\Gamma}) = 0 . \tag{5}$$

Using Eq. (5) in Eq. (2), and integrating by parts (the boundary terms vanish, either because of periodicity or because the distribution vanishes at the boundaries), we have

$$\frac{d}{dt}\langle B(t)\rangle = \int d\Gamma\, B\frac{\partial f}{\partial t}$$

$$= -\int d\Gamma\, B\frac{\partial}{\partial\Gamma}\cdot(f\dot\Gamma)$$

$$= \int d\Gamma\, f\dot\Gamma\cdot\frac{\partial B}{\partial\Gamma}$$

$$= \int d\Gamma\, f\dot B$$

$$= \langle\dot B(t)\rangle\ ; \tag{6}$$

that is, the operations d/dt and $\langle\,\cdots\,\rangle$ commute if f obeys the Liouville equation. Had we replaced the right-hand side of Eq. (5) with a diffusive term, such as

$$\frac{\partial f}{\partial t}+\frac{\partial}{\partial\Gamma}\cdot(f\dot\Gamma)=D\frac{\partial}{\partial\Gamma}\cdot\left[\frac{\partial f}{\partial\Gamma}\right]\ , \tag{7}$$

then we would have had

$$\frac{d}{dt}\langle B(t)\rangle = \int d\Gamma\, B\frac{\partial f}{\partial t}$$

$$= \int d\Gamma\, B\left[-\frac{\partial}{\partial\Gamma}\cdot(f\dot\Gamma)+D\frac{\partial}{\partial\Gamma}\cdot\left[\frac{\partial f}{\partial\Gamma}\right]\right]$$

$$= \int d\Gamma\, f\dot\Gamma\cdot\frac{\partial B}{\partial\Gamma}-D\int d\Gamma\, \frac{\partial B}{\partial\Gamma}\cdot\frac{\partial f}{\partial\Gamma}$$

$$= \int d\Gamma\, f\dot B+D\int d\Gamma\, f\frac{\partial}{\partial\Gamma}\cdot\frac{\partial B}{\partial\Gamma}$$

$$= \langle\dot B(t)\rangle+D\left\langle\frac{\partial}{\partial\Gamma}\cdot\frac{\partial B}{\partial\Gamma}\right\rangle\ ; \tag{8}$$

that is, for $D\neq0$, d/dt and $\langle\,\cdots\,\rangle$ do *not* commute. Therefore, Eq. (7) is obviously *not* a viable replacement for the Liouville equation [Eq. (5)].

Now consider feedback-thermostatted equations of motion where the momentum equation of motion $\dot p$ includes a damping term $-\nu\zeta p$; ν is the rate of thermostatting (a constant) and ζ is the dimensionless friction coefficient, an extra degree of freedom in phase space. We propose the following candidate to replace the Liouville equation—for lack of a better name, we will call it the mixing Liouville equation:

$$\frac{\partial f}{\partial t}+\frac{\partial}{\partial\Gamma}\cdot(f\dot\Gamma)=\lambda\frac{\partial^2(f-f_0)}{\partial\zeta^2}\ , \tag{9}$$

where the time-irreversible mixing or diffusion in phase space is primarily normal to the usual (q,p) subspace for the system of particles, driven by the difference in the distribution function from its equilibrium value, $\Delta f=f-f_0$, at the rate λ, which is to be regarded as empirical and perhaps equal to the Lyapunov exponent.[6] (It is reasonable to suppose that λ is a function of the thermostatting rate ν as well as the external applied field ϵ.) The macroscopic equation of motion for B then reads

$$\frac{d}{dt}\langle B(t)\rangle = \int d\Gamma\, B\frac{\partial f}{\partial t}=\int d\Gamma\, B\left[-\frac{\partial}{\partial\Gamma}\cdot(f\dot\Gamma)\right.$$

$$\left.+\lambda\frac{\partial^2\Delta f}{\partial\zeta^2}\right]$$

$$= \int d\Gamma\, f\dot B+\lambda\int d\Gamma\, \Delta f\frac{\partial^2 B}{\partial\zeta^2}$$

$$= \langle\dot B(t)\rangle\ , \tag{10}$$

when integrations by parts have been performed (boundary terms in ζ vanish because the distribution and its ζ derivatives vanish at $\zeta=+\infty$ or $-\infty$). The term proportional to λ vanishes because we are primarily interested in observables B either that depend only on the (q,p) subspace, in which case $\partial B/\partial\zeta=0$, or whose dependence on ζ is at most quadratic, in which case $\partial^2 B/\partial\zeta^2$ is a constant multiplying $\int d\Gamma\, \Delta f=0$ since both f and f_0 are normalized. Hence, the operations d/dt and $\langle\,\cdots\,\rangle$ commute if the Liouville equation is replaced by Eq. (9).

It is important to note at this point that the Heisenberg-Schrödinger equivalence of classical response theory[1] is preserved by the foregoing commutativity of d/dt and $\langle\,\cdots\,\rangle$. However, the f propagator for the distribution function is purely Liouvillean; that is, because the non-Liouville part of Eq. (9) does not affect the macroscopic equation of motion for the observable, f *appears* to obey the Liouville equation—at least as far as ensemble (equilibrium or nonequilibrium) averages of dynamical observables such as $B(\Gamma)$ are concerned. This Heisenberg-Schrödinger (or Lagrangian-Eulerian) equivalence does *not* apply to the entropy, which cannot be formulated as a local property of a single phase-space trajectory. Entropy is intrinsically a Schrödinger (Eulerian) property depending on an entire ensemble of trajectories. Furthermore, the entropy is a nonlinear functional of f, while ensemble averages such as Eq. (1) are linear. Thus, any replacement for the Liouville equation must also satisfy a second criterion, namely, that it describes the time evolution of the entropy correctly.

The entropy of a nonequilibrium system is given by the following expression [provided that the "equation of motion" of the distribution function is the mixing Liouville equation given by Eq. (9)]:

$$S(t)=S_0-k\int d\Gamma\, f(\Gamma,t)\ln[f(\Gamma,t)/f_0(\Gamma)]\ , \tag{11}$$

where the equilibrium entropy is $S_0=-k\int d\Gamma\, f_0\ln f_0$. The nonequilibrium entropy, relative to the equilibrium distribution (the so-called "distribution of complete ignorance"[7]), is

$$\Delta S(t)=S(t)-S_0=-k\int d\Gamma\, f\ln(f/f_0)\ , \tag{12}$$

which is the difference of the information-theory "entropy"

$$S_{\text{info}}(t)=-k\int d\Gamma\, f\ln f\ , \tag{13}$$

and the Zubarev constant-free-energy "entropy"

$$S_{\text{Zub}}(t)=-k\int d\Gamma\, f\ln f_0\ . \tag{14}$$

(The only reason to discuss these separately is historical[8] as well as for convenience.) The macroscopic equation of motion for the entropy is $\dot{S}(t) = \dot{S}_{\text{info}}(t) + \dot{S}_{\text{Zub}}(t)$. First of all,

$$\dot{S}_{\text{info}}/k = -\frac{d}{dt} \int d\Gamma\, f \ln f$$

$$= -\int d\Gamma \frac{\partial}{\partial t}(f \ln f)$$

$$= -\int d\Gamma (\ln f + 1)\frac{\partial f}{\partial t}$$

$$= -\int d\Gamma \ln f \frac{\partial f}{\partial t} , \qquad (15)$$

since $\int d\Gamma\, \partial f/\partial t = 0$ from the normalization of f. From Eq. (9), we have

$$\dot{S}_{\text{info}}/k = -\int d\Gamma \ln f \left[-\frac{\partial}{\partial \Gamma} \cdot (f\dot{\Gamma}) + \lambda \frac{\partial^2 \Delta f}{\partial \zeta^2} \right]$$

$$= -\int d\Gamma \left[f\dot{\Gamma} \cdot \frac{\partial \ln f}{\partial \Gamma} + \lambda \ln f \frac{\partial^2 \Delta f}{\partial \zeta^2} \right]$$

$$= -\int d\Gamma \left[\dot{\Gamma} \cdot \frac{\partial f}{\partial \Gamma} + \lambda \ln f \frac{\partial^2 \Delta f}{\partial \zeta^2} \right]$$

$$= \int d\Gamma\, f \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} - \lambda \int d\Gamma \ln f \frac{\partial^2 \Delta f}{\partial \zeta^2}$$

$$= \langle \Omega \rangle - \lambda \int d\Gamma \ln f \frac{\partial^2 \Delta f}{\partial \zeta^2} , \qquad (15a)$$

where $\Omega = (\partial/\partial \Gamma)\cdot\dot{\Gamma}$ is the phase-space compressibility (see Appendix A). Secondly, we have

$$\dot{S}_{\text{Zub}}/k = -\frac{d}{dt} \int d\Gamma\, f \ln f_0$$

$$= -\int d\Gamma \ln f_0 \frac{\partial f}{\partial t}$$

$$= -\int d\Gamma \ln f_0 \left[-\frac{\partial}{\partial \Gamma} \cdot (f\dot{\Gamma}) + \lambda \frac{\partial^2 \Delta f}{\partial \zeta^2} \right]$$

$$= -\int d\Gamma \left[f\dot{\Gamma} \cdot \frac{\partial \ln f_0}{\partial \Gamma} + \lambda \ln f_0 \frac{\partial^2 \Delta f}{\partial \zeta^2} \right]$$

$$= -\int d\Gamma \left[f\frac{1}{f_0}\dot{\Gamma} \cdot \frac{\partial f_0}{\partial \Gamma} + \lambda \ln f_0 \frac{\partial^2 \Delta F}{\partial \zeta^2} \right]$$

$$= \beta \int d\Gamma\, f\dot{H} - \lambda \int d\Gamma \ln f_0 \frac{\partial^2 \Delta f}{\partial \zeta^2}$$

$$= \beta \dot{E} - \lambda \int d\Gamma \ln f_0 \frac{\partial^2 \Delta f}{\partial \zeta^2} , \qquad (16)$$

where we have used the fact that for the canonical ensemble $\dot{\Gamma}\cdot(\partial f_0/\partial \Gamma) = -\beta \dot{H}f_0$ (see Appendix A for details). Also in Appendix A, we show that $\beta \dot{E} = \langle \Omega \rangle - \beta \dot{W}$, where $\dot{W}$ is the power *out* of the system of particles (when external forces drive the system, $\dot{W} < 0$). Combining Eqs. (15) and (16), we get

$$\dot{S}/k = \beta \dot{W} - \lambda \int d\Gamma \ln \left[\frac{f}{f_0} \right] \frac{\partial^2 \Delta f}{\partial \zeta^2}$$

$$= \beta \dot{W} + \lambda \int d\Gamma \frac{\partial \ln(f/f_0)}{\partial \zeta} \frac{\partial \Delta f}{\partial \zeta}$$

$$= \beta \dot{W} + \lambda \int d\Gamma \frac{1}{f} \left[\left[\frac{\partial \Delta f}{\partial \zeta} \right]^2 + g\zeta \frac{\partial \Delta f}{\partial \zeta} \Delta f \right] , \qquad (17)$$

where we have used the fact that $\partial f_0/\partial \zeta = -g\zeta f_0$ (see Appendix A).

There are four cases for which we would like to test the entropy equation of motion [Eq. (17)]: (1) equilibrium, (2) adiabatic (unthermostatted) external driving, (3) the thermostatted nonequilibrium steady state, and (4) equilibration (heating or cooling) from a nonequilibrium initial state.

(1) *Equilibrium.* $\dot{W} \equiv 0$ and $f = f_0$; hence, $\dot{S} = 0$ and $S = S_0$.

(2) *Adiabatic external driving.* Since $v = 0$, $\dot{Q} \equiv 0$ and $\Omega \equiv 0$; hence, $\dot{E}_0 = -\dot{W} = \dot{E} > 0$. Also, the equations of motion leaving ζ unchanged and do not depend on the initial values of ζ; hence the distribution function f cannot depend on ζ. We can therefore write $\partial \Delta f/\partial \zeta = -\partial f_0/\partial \zeta = g\zeta f_0$, and we find that

$$\dot{S}_{\text{ad}}/k = -\beta \dot{E}_0 + g^2 \lambda \int d\Gamma\, f_0 \zeta^2$$

$$= -\beta \dot{E}_0 + g\lambda , \qquad (18)$$

since $\langle \zeta^2 \rangle_0 = \int d\Gamma\, f_0 \zeta^2 = 1/g$. Consider now two sub-cases.

(a) $\dot{E}_0/E_0 \ll \nu_{\text{collision}}$ (slow, reversible process; i.e., the work on the system is done much more slowly than any natural microscopic collision rate). In this case, we expect $\dot{S}_{\text{ad}} \equiv 0$, that is, an isentropic increase in the temperature; the non-Liouville isentropic rate λ_s is therefore

$$\lambda_s = \dot{E}_0/gkT > 0 . \qquad (18a)$$

(b) $\dot{E}_0/E_0 \gg \nu_{\text{collision}}$ (fast, irreversible shock process; i.e., the work on the system is done too fast for the system to adjust). Here, the entropy of the system increases irreversibly, so that all we can say about λ is that $\lambda > \lambda_s > 0$:

$$\lambda > \dot{E}_0/gkT . \qquad (18b)$$

(3) *Thermostatted nonequilibrium steady state.* In the linear-response regime, $\Delta f \sim \epsilon^2$ (small); from Appendix A, we see that $\dot{W}_{\text{ss}} \sim \epsilon^2$ and that $\langle \zeta \rangle_{\text{ss}} \sim \dot{W} \sim \epsilon^2$. Thus $(\partial \Delta f/\partial \zeta)^2 \sim \epsilon^2$, while $g\zeta(\partial \Delta f/\partial \zeta)\Delta f \sim \epsilon^4$ can be ignored in Eq. (17). Since $\dot{S}_{\text{ss}} = 0$,

$$\lambda_{\text{ss}} = -\beta \dot{W}_{\text{ss}} \bigg/ \int d\Gamma \frac{1}{f_{\text{ss}}} \left[\frac{\partial \Delta f_{\text{ss}}}{\partial \zeta} \right]^2 > 0 , \qquad (19)$$

because $\dot{W}_{\text{ss}} < 0$. Notice one very important thing: if the Liouville equation were entirely correct ($\lambda \equiv 0$), then $T\dot{S}_{\text{ss}} = \dot{W}_{\text{ss}} < 0$, which is nonsensical unless the expression for the entropy [Eq. (11)] were wrong.

(4) *Thermal equilibration.* Imagine that an ensemble is started out with a distribution characterized by tempera-

ture T_0 different from the thermostat (for convenience, we let the thermostat temperature $T = 1$, the mass of each particle $m = 1$, and Boltzmann's constant $k = 1$). We also imagine that the initial distribution $f \sim f_0^{1/T_0}$, so that $\langle p^2 \rangle_0 = 1 = g \langle \zeta^2 \rangle_0$ corresponds to expectations from f_0 at equilibrium, while $\langle p^2(0) \rangle = T_0 = g \langle \zeta^2(0) \rangle$ corresponds to f at $t = 0$. Consequently, with this simplification, at early times we have

$$\frac{\partial f}{\partial \zeta} = \frac{1}{T_0} f_0^{(1/T_0 - 1)} \frac{\partial f_0}{\partial \zeta} = -\frac{g\zeta}{T_0} f \ . \tag{20}$$

Since $\dot{W} \equiv 0$, Eq. (17) becomes, for early times,

$$\dot{S} = g\lambda \left[\frac{1}{T_0} - 1 \right] \int d\Gamma \, g\zeta^2 \left[\frac{f}{T_0} - f_0 \right]$$

$$= g\lambda \left[\frac{1}{T_0} - 1 \right] \left[\frac{g}{T_0} \langle \zeta^2 \rangle - g \langle \zeta^2 \rangle_0 \right]$$

$$= g\lambda \left[\frac{1}{T_0} - 1 \right] \left[\frac{g}{T_0} \langle \zeta^2 \rangle - 1 \right] \ . \tag{21}$$

At $t = 0$, $\dot{S} = 0$; but for times on the order of the thermo-statting relaxation time, $\tau = 1/\nu$, $g \langle \zeta^2 \rangle / T_0$ departs from unity according to whether $T_0 < 1$ (heating) or $T_0 > 1$ (cooling):

(a) $T_0 < 1$ ("Big Bang"). In this case, the thermostat heats the system up, so that $\langle \zeta \rangle$ decreases from its initial value of zero. For times $t \sim \tau$, $g \langle \zeta^2 \rangle / T_0 \to 1 + \delta$, where $\delta > 0$, so that if $T_0 \ll 1$,

$$\dot{S} \to \delta g \lambda / T_0 > 0 \ . \tag{21a}$$

(b) $T_0 > 1$ ("Big Shrink"). Here, the thermostat cools the system down, so that $\langle \zeta \rangle$ increases from its initial value of zero. At first, for times $t \sim \tau$, $g \langle \zeta^2 \rangle / T_0 \to 1 + \delta$, where $\delta > 0$, so that if $T_0 \gg 1$, $\dot{S} \to -\delta g \lambda < 0$. At longer times, $g \langle \zeta^2 \rangle / T_0 \to 1 - \delta$ as $g \langle \zeta^2 \rangle$ heads down toward the equilibrium value of unity; then

$$\dot{S} \to \delta g \lambda > 0 \ . \tag{21b}$$

In both cases of undriven thermostatted equilibration, the entropy is seen to rise from a lower initial nonequilibrium value toward equilibrium (see Figs. 1 and 2). Since the internal energy of the system rises toward equilibrium for the "Big Bang" experiment, so that for early times $(t \sim \tau)$

$$\dot{E}_0 = -2\nu \langle K\zeta \rangle \approx -g\nu T_0 \langle \zeta \rangle > 0 \ , \tag{22}$$

it is therefore reasonable to expect the entropy to rise fast enough that the free energy, $A = E_0 - TS$, drops toward equilibrium, where it should have a minimum value. Indeed, from Eq. (21a), we see that this expectation is reasonable: both δ and T_0 are small, so their ratio is of order unity, while in the expression for $\dot{E}_0$, both $\langle \zeta \rangle$ and T_0 are small, so their product is smaller yet. For the "Big Shrink" experiment, both $\dot{E}_0$ and $-\dot{S}$ are negative (after initial transients in S), so that the free energy must approach a minimum at equilibrium.

In contrast to these predictions of the mixing Liouville equation, if the Liouville equation were correct ($\lambda \equiv 0$),
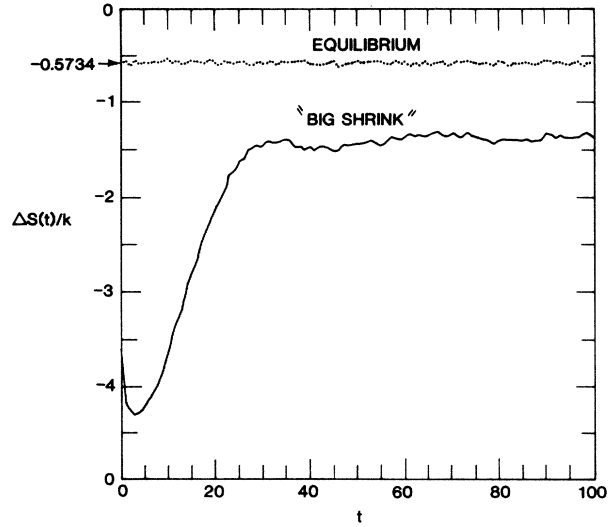

FIG. 1. Entropy relative to equilibrium for the "Big Shrink" $(T_0 > T)$ experiment for the Nosé-Hoover oscillator ensemble $(N = 1000, \nu = T = 1, T_0 = 100)$. Note that the response time is $15-20 \sim 1/\lambda$. The Lyapunov exponent (largest positive) is $\lambda \sim 0.04$ (Ref. 9). Equilibrium is not perfectly attained because of the nonergodicity of the Nosé-Hoover oscillator, as discussed in Ref. 2. (An equilibrium realization is shown as a a dashed line.) See Appendix B for computational details.

Eq. (17) would predict that S is a constant of the motion $(\dot{S} = 0)$ for both cases. It is remarkable that the early behavior of the entropy is predicted so well by the mixing Liouville equation, including such detailed features as the initial drop in entropy and its gradual rise for the "Big Bang" experiment $(T_0 < 1)$. (As is discussed in more de-
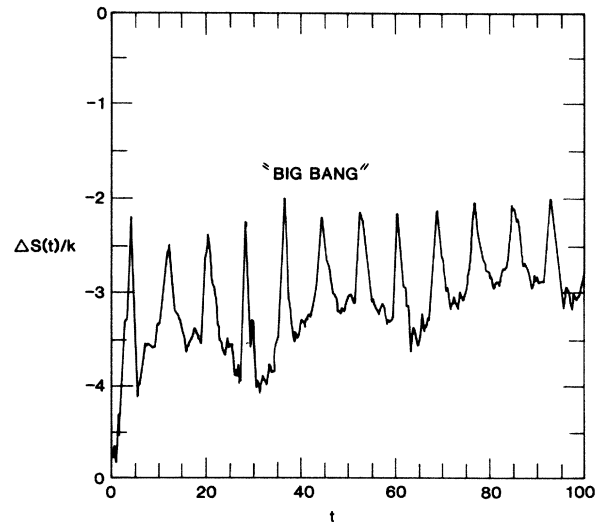

FIG. 2. Entropy relative to equilibrium for the "Big Bang" $(T_0 < T)$ experiment for the Nosé-Hoover oscillator ensemble $(N = 1000, \nu = T = 1, T_0 = 0.01)$. Note that the response time is $\sim 2 \ll 1/\lambda$ [compare Eqs. (21a) and (21b) in the text]; $\lambda \sim 0.04$ (Ref. 9). As discussed in Ref. 2, the ringing of period ~ 8 can be related to the size of the second-law attractor for initial conditions near the origin.

tail in Ref. 2, the ringing in this experiment is the result of a flushing motion down the ζ axis whose time constant is governed by the thermostatting time and the size of the second-law attractor.) In both cases, λ is close to the largest Lyapunov exponent for the Nosé-Hoover oscillator.[9] We conclude from these results that the mixing Liouville equation [Eq. (9)] is a valid replacement for the Liouville equation.

III. CONCLUSIONS

We have shown that the mixing Liouville equation

$$\dot{f} + f\Omega = \lambda \partial^2 \Delta f / \partial \zeta^2$$

correctly describes the time evolution of the distribution function under a wide range of experimental conditions for thermostatted dynamical systems. The time evolution of the entropy provides a critical test of this equation, since we show that the ensemble averages of typical mechanical observables are also correctly described by the Liouvillle equation ($\lambda \equiv 0$). We have presented numerical evidence for the validity of the mixing Liouville equation, whereby the entropy relative to equilibrium

$$\Delta S = -k \int d\Gamma f \ln(f/f_0)$$

increases with thermal equilibration (either thermostatted heating or cooling) in agreement with experiment; the usual Liouville equation, on the other hand, predicts that ΔS is a constant of the motion, in disagreement with the numerical results. The mixing Liouville equation is irreversible in time for nonequilibrium initial and boundary conditions, particularly for times greater than $1/\lambda$; at equilibrium ($\Delta f = 0$), it is reversible. Diffusion or mixing in phase space is in the direction of the thermostatting friction coefficient ζ, orthogonal to the (q,p) subspace of the particles, such that the distribution is driven toward equilibrium on a time scale of $1/\lambda$.

ACKNOWLEDGMENTS

APPENDIX A

For thermostatted equations of motion with external driving $\dot{\epsilon}(t)$, e.g., shear flow at shear rate $\dot{\epsilon} = \partial u_x / \partial y$, $\Gamma(\Gamma,t) = (\dot{q}, \dot{p}, \dot{\zeta})$:

$$\dot{q} = p/m + \dot{\epsilon} \hat{Q} ,$$

$$\dot{p} = F - \dot{\epsilon} \hat{P} - \nu \zeta p ,$$

$$\dot{\zeta} = \nu(K/K_0 - 1) , \tag{A1}$$

where $\hat{Q}$ and $\hat{P}$ are functions of coordinates q and momenta p such that $\Sigma(\partial/\partial q)\hat{Q} = \Sigma(\partial/\partial p)\hat{P}$, $F = -(\partial \Phi/\partial q)$

is the force due to the potential energy of particle interactions Φ, and the kinetic energy is $K = \Sigma p^2/2m$ for no center-off-mass motion. With the characteristic rate ν, the thermostat keeps the time-averaged, as well as ensemble-averaged value of K equal to $K_0 = gkT/2$, where g is the number of degrees of freedom and T is the temperature of the thermal reservoir. The phase-space compressibility can be shown to be

$$\Omega = \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} = -g\nu\zeta . \tag{A2}$$

The internal energy is $H_0 = K + \Phi$, so that its rate of change is

$$\dot{H}_0 = \dot{\Gamma} \cdot \frac{\partial H_0}{\partial \Gamma} = -2\nu K \zeta - J\dot{\epsilon} , \tag{A3}$$

where the dissipative energy flux function is

$$J = \Sigma(p\hat{P}/m + F\hat{Q}) . \tag{A4}$$

With the definition $E_0 = \langle H_0 \rangle$, and since $\dot{E}_0 = \langle \dot{H}_0 \rangle$ by virtue of Eq. (10) (that is, $\partial H_0/\partial \zeta = 0$), we have

$$\dot{E}_0 = \dot{Q} - \dot{W} , \tag{A5}$$

where the heat flow *into* the system of particles is

$$\dot{Q} = -2\nu\langle K\zeta \rangle \tag{A6}$$

and the rate at which the system of particles does work on the external world, or the power *out* of the system, is

$$\dot{W} = \langle J \rangle \dot{\epsilon} . \tag{A7}$$

A so-called "adiabatic" system, where no heat flows into or out of the system ($\dot{Q} = 0$) can be achieved by turning the thermostat off, that is, by setting the thermostatting rate to zero, $\nu = 0$. In that case, the phase-space compressibility is also zero [Eq. (A2)]. With the thermostat turned on, $\nu > 0$, a steady state can be achieved, where the work done on the system by a steady externally applied field $\dot{\epsilon}(t) = \dot{\epsilon}$ is balanced by the heat extracted from the system to keep it at a constant temperature: $\dot{E}_0^{ss} = 0$, $\dot{Q}_{ss} = \dot{W}_{ss} < 0$. (For example, if the field is a steady shear rate applied to a fluid, $\dot{\epsilon} = \partial u_x/\partial y$, then $\langle J \rangle_{ss} = \langle P_{yx} \rangle_{ss} V = -\eta V\dot{\epsilon}$, where P_{yx} is the shear component of the pressure tensor, V is the constant volume of the system of N particles, and η is the linear shear viscosity. At the steady state, $\dot{Q}_{ss} = -2\nu\langle K\zeta \rangle_{ss} = -2\nu K_0\langle \zeta \rangle_{ss} = \dot{W}_{ss} = -\eta V\dot{\epsilon}^2 < 0$; hence, the steady-state value of the thermostatting coefficient is positive: $\langle \zeta \rangle_{ss} = \eta V\dot{\epsilon}^2/gkT\nu$. The phase-space compressibility of the steady state is therefore negative: $\langle \Omega \rangle_{ss} = -g\nu\langle \zeta \rangle_{ss}$.)

The equilibrium distribution function for the canonical ensemble f_0 can be shown to be a stationary solution of both the Liouville equation [Eq. (5)] and the mixing Liouville equation [Eq. (9)] when the thermostatted ($\nu > 0$) equations of motion [Eq. (A1)] govern the system and the external driving field is set to zero: $\dot{\epsilon}(t) = 0$. We define the total energy of the system, including a contribution from the thermostat, to be $H = H_0 + gkT\zeta^2/2$. Then, the equilibrium distribution for temperature T ($\beta = 1/kT$) is

$$f_0(\Gamma) = Z^{-1}(\beta)\exp[-\beta H(\Gamma)] \ , \qquad (A8)$$

where $Z(\beta) = \int d\Gamma \exp(-\beta H)$ is the normalization (partition function). Note that f_0 depends on the phase Γ only through the total energy H, so that, by the chain rule,

$$\dot{\Gamma} \cdot \frac{\partial f_0}{\partial \Gamma} = \dot{\Gamma} \cdot \frac{\partial H}{\partial \Gamma} \frac{\partial f_0}{\partial H} = -\beta \dot{H} f_0 \ . \qquad (A9)$$

The rate of change of the total energy is

$$\dot{H} = \dot{\Gamma} \cdot \frac{\partial H}{\partial \Gamma} = \dot{H}_0 + gkT\zeta\dot{\zeta} = -2\nu K_0 \zeta - J\dot{\epsilon}$$

$$= kT\Omega - J\dot{\epsilon} \ . \qquad (A10)$$

At equilibrium, $\dot{\epsilon} = 0$ and $\beta \dot{H} = \Omega$; inserting $f = f_0$ into Eq. (9) and noting that $\partial f_0/\partial t = 0$, we see that f_0 is a stationary, nonzero solution for both Eq. (5) and Eq. (9):

$$\frac{\partial f_0}{\partial t} + \dot{\Gamma} \cdot \frac{\partial f_0}{\partial \Gamma} + f_0 \frac{\partial}{\partial \Gamma} \cdot \dot{\Gamma} = f_0(-\beta \dot{H} + \Omega) = 0 \ .$$

Another property of f_0 that we use in the main text is due to the quadratic dependence of H upon ζ:

$$\frac{\partial f_0}{\partial \zeta} = -g\zeta f_0 \ . \qquad (A11)$$

Finally, with the definition $E = \langle H \rangle$, and since $\dot{E} = \langle \dot{H} \rangle$ by virtue of Eq. (10) (that is, $\partial^2 H/\partial \zeta^2 = gkT$), we obtain from Eq. (A10)

$$\dot{E} = kT\langle \Omega \rangle - \dot{W} \ . \qquad (A12)$$

APPENDIX B

In computing ensemble averages of ordinary mechanical observables B, such as energy, for a finite number of ensemble members N, we use the Heisenberg picture for convenience, basing the individual weights on the initial distribution $f(\Gamma,0)$:

$$\langle B(t) \rangle = \int d\Gamma f(\Gamma,0)B(U(t)\Gamma)$$

$$\approx \sum_{i=1}^{N} \Delta\Gamma_i f(\Gamma_i,0)B(U(t)\Gamma_i)$$

$$= \frac{1}{N} \sum_{i=1}^{N} B_i(t) \ , \qquad (B1)$$

where $B_i(t) = B(U(t)\Gamma_i)$ is the observable evaluated along trajectory i. The phase-space propagator $U(t)$ moves the phase along the trajectory from its $t = 0$ value, Γ. The trajectory weights are chosen to be equal, so that $\Delta\Gamma_i f(\Gamma_i,0) = 1/N$ determines the volume elements which guarantee that the distribution is normalized.

In the Schrödinger picture, which is equivalent to the Heisenberg view for ordinary observables, we fix the cells in phase space so that on the average at equilibrium, there is one ensemble element per cell: $\Delta\Gamma_i f_0(\Gamma_i) = 1/N$. We can easily do this for the Nosé-Hoover oscillator ($T = 1$), since $\Gamma = (q,p,\zeta)$ and f_0 is a simple Gaussian:

$$f_0(q,p,\zeta) = (2\pi)^{-3/2}\exp[-\tfrac{1}{2}(q^2+p^2+\zeta^2)] \ , \qquad (B2)$$

so that for ensemble element n ($= 1,2,\ldots,N$) the cell number m in phase space is obtained from

$$i = \mathrm{int}\{\tfrac{1}{2}N^{1/3}[1+\mathrm{erf}(2^{-1/2}q_n)]+1\} \ , \qquad (B3)$$

with j and k determined similarly for p_n and ζ_n;

$$m = i + N^{1/3}[j-1+N^{1/3}(k-1)] \qquad (B3a)$$

is a number between 1 and N. In Eq. (B3), the normal error function is

$$\mathrm{erf}(x) = 2\pi^{-1/2} \int_0^x dy \exp(-y^2)$$

and "int" stands for the integer part.

The Schrödinger equivalent of Eq. (B1) is then

$$\langle B(t) \rangle = \int d\Gamma f(\Gamma,t)B(\Gamma)$$

$$\approx \sum_{i=1}^{N} \Delta\Gamma_i f(\Gamma_i,t)B(\Gamma_i)$$

$$= \frac{1}{N} \sum_{i=1}^{N} n_i(t)B_i \ , \qquad (B4)$$

where $B_i = B(\Gamma_i)$ is the value of the observable in cell i and $n_i(t) = f(\Gamma_i,t)/f_0(\Gamma_i)$ is the instantaneous number of ensemble members in the cell relative to the average of one per cell at equilibrium.

For ordinary observables, the Schrödinger (Eulerian) picture of Eq. (B4) is not as smooth as Eq. (B1), the Heisenberg (Lagrangian) version. For the entropy as defined in Eq. (11), we have no choice but to use the Schrödinger picture:

$$\Delta S(t)/k = -\int d\Gamma f(\Gamma,t)\ln\left[\frac{f(\Gamma,t)}{f_0(\Gamma)}\right]$$

$$\approx -\sum_{i=1}^{N} \Delta\Gamma_i f(\Gamma_i,t)\ln\left[\frac{f(\Gamma_i,t)}{f_0(\Gamma_i)}\right]$$

$$= -\frac{1}{N} \sum_{i=1}^{N} n_i(t)\ln n_i(t) \ . \qquad (B5)$$

This numerical realization of the entropy relative to its equilibrium value as given by Eq. (B5) is shown as the upper curve of Fig. 1 for an equilibrium ensemble. Note that ΔS is fluctuating about a constant nonzero value. This paradoxical result is immediately explained by the discrete nature of Eq. (B5), which can be rewritten in terms of the probability of finding k elements of the ensemble in a given cell, p_k ($k = 0,1,2,\ldots,N$):

$$\Delta S/k = -\sum_{k=0}^{N} p_k k \ln k$$

$$\approx -e^{-1}\left[\frac{2\ln 2}{2!} + \frac{3\ln 3}{3!} + \cdots\right]$$

$$= -0.5734 \ , \qquad (B6)$$

where we have substituted the (normalized) Poisson distribution $e^{-1}/k!$ for p_k at equilibrium;[10] this result is consistent with an average of one element per cell at equilibrium as shown by

$$\langle n \rangle = \sum_{k=0}^{N} p_k k$$

$$\approx e^{-1} \left[\frac{0}{0!} + \frac{1}{1!} + \frac{2}{2!} + \frac{3}{3!} + \cdots \right]$$

$$= 1 \ . \tag{B7}$$

The Poisson distribution approximation becomes exact as the number of ensemble members goes to infinity. Notice that $N = 1000$ in Fig. 1 (dashed curve for equilibrium) is already an excellent approximation to the infinite limit.

This is *not* the same as the thermodynamic limit however, since each ensemble member represents a unique initial condition and resulting trajectory for *one* one-dimensional thermostatted harmonic oscillator.

Finally, we point out that the qualitative increase of ΔS from a low initial value toward a maximum at equilibrium, as shown in Figs. 1 and 2 for the "Big Shrink" ($T_0 > 1$) and "Big Bang" ($T_0 < 1$) experiments, is *not* an artifact of the discrete evaluation. The approach toward the second-law attractor is also qualitatively observed in the apparently irreversible change in the size of the cloud of ensemble members.

---

[1]B. L. Holian and D. J. Evans, J. Chem. Phys. **83**, 3560 (1985); B. L. Holian, *ibid.* **84**, 1762 (1986).

[2]B. L. Holian and W. G. Hoover, preceding paper, Phys. Rev. A **34**, 4229 (1986).

[3]S. Nosé, J. Chem. Phys. **81**, 511 (1984); W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).

[4]D. J. Evans and B. L. Holian, J. Chem. Phys. **83**, 4069 (1985).

[5]B. L. Holian, J. Chem. Phys. **84**, 3138 (1986).

[6]The relationship of the largest Lyapunov exponent to the Kolmogorov "entropy" (rate of divergence of trajectories) is discussed by G. Bennetin, L. Galgani, and J.-M. Strelcyn, Phys. Rev. A **14**, 2338 (1976).

[7]E. T. Jaynes: *Papers on Probability, Statistics, and Statistical Physics,* edited by R. D. Rosenkrantz (Reidel, Dordrecht, Holland, 1983), pp. 59, 60, and 166ff; see also J. D. Ramshaw, Phys. Lett. **A116**, 110 (1986); and T. Feldmann, B. Andresen, A. Qi, and P. Salamon, J. Chem. Phys. **83**, 5849 (1985).

[8]B. L. Holian, Phys. Rev. A **33**, 1152 (1986).

[9]H. A. Posch, W. G. Hoover, and F. J. Vesely, Phys. Rev. A **33**, 4253 (1986).

[10]W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1968), Vol. I, Chap. IV.