PHYSICAL REVIEW A VOLUME 39, NUMBER 10 MAY 15, 1989

Nonlinear-response theory for time-independent fields: Consequences of the fractal nonequilibrium distribution function

Brad Lee Holian
Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Giovanni Ciccotti
Dipartimento di Fisica, Università degli Studi di Roma "La Sapienza," I-00185 Roma, Italy

William G. Hoover
University of California at Davis–Livermore and Lawrence Livermore National Laboratory, Livermore, California 94550

Bill Moran
Lawrence Livermore National Laboratory, Livermore, California 94550

Harald A. Posch
Institut für Experimentalphysik der Universität Wien, A-1090 Wien, Austria
(Received 3 October 1988)

Classical nonlinear-response theory is applied to an ensemble of experiments performed under the same external field. If the field is time independent, the formalism simplifies, and we can derive the "Kawasaki expression" for the distribution function, as well as the transient correlation-function expression for the observable response. Recent computer simulations have shed new light on our understanding of the steady-state nonequilibrium distribution function: the fractal nature of the distribution causes us to reevaluate the usual mathematical tools used to describe nonequilibrium processes.

I. INTRODUCTION

The purpose of classical nonlinear-response theory is to provide a mathematical framework for the measurement of nonequilibrium, time-dependent processes.[1] We imagine performing a large number, or ensemble, of identical experiments on a sample, such as a fluid sandwiched between hot and cold walls. The elements of the ensemble differ from each other only in their initial conditions; otherwise, for time $t>0$, each is subjected to the same boundary conditions in the form of a time-dependent external driving force $X(t)$ turned on at $t=0$. The transient response of each system is monitored, and the experimental time-histories for the ensemble are then averaged together to refine the statistical precision of the measurement.

The set of initial conditions for the ensemble of experiments is most conveniently chosen from the equilibrium canonical distribution f_0. Each member of the ensemble is an N-particle system (the sample) represented at $t=0$ by a point in phase space, which thenceforth moves along a trajectory independently of all others in the ensemble. The dimensionality of the classical-mechanical phase space, including all coordinates q and momenta p is $2N$ times the number of spatial (Cartesian) dimensions. It is reasonable to imagine that a continuous function of time t at each point Γ in phase space, that is, the nonequilibrium distribution function $f(\Gamma,t)$, might be capable of describing the nonequilibrium dynamics of this ensemble—"a swarm of blind flies." The "swarm" is a peculiar ideal gas: trajectories do not "see" each other at all, so there are no conservation equations for "momentum" or "energy" of the ensemble, only conservation of "mass." Each trajectory ("fly"), initially located in a volume element $d\Gamma$ at phase point Γ, carries its own share of the weight of the ensemble ("swarm"), as measured at time zero by $f(\Gamma,0)d\Gamma=f_0(\Gamma)d\Gamma$, with the total always summing to unity: $\int d\Gamma\, f(\Gamma,t)=1$.

This apparently satisfactory physical picture of an ensemble, whose dynamics is described using mathematics appropriate to a regular, continuous distribution $f(\Gamma,t)$, is confronted by a paradox first noted by Gibbs.[2] That is, suppose one performs a relaxation experiment on an ensemble whose initial distribution $f(\Gamma,0)$ is perturbed from equilibrium. If the members of the ensemble are then isolated from the rest of the world and allowed to reach equilibrium, the Gibbs *mathematical* entropy for the ensemble remains a constant during the relaxation process, rather than increasing as the ensemble approaches equilibrium. For example, imagine that each member of the ensemble is a perfect crystal, where the atoms are initially given momenta randomly selected from a Maxwell-Boltzmann distribution for a temperature twice that of melting. The question is, will each such isolated system melt, and if so, will the *physical* entropy increase as equilibrium is approached? If the answer should reasonably be "yes" to both parts, then why does the mathematics predict no change in the entropy? For the case of a steady external force that *drives* the system away from equilibrium, with a thermostat that keeps the temperature constant, the corresponding paradox is that the entropy continues to drop *forever,* even at the steady state.

On the other hand, in spite of this entropy conundrum, a great deal of insight into nonequilibrium processes at the atomistic level has been gained from the simulation on the computer of the motion of atoms, using the method of molecular dynamics (MD).[3] In these computer experiments, the observable responses to external driving forces have been measured and found to be consistent with the second law of thermodynamics. That is, transport coefficients (ratio of response to driving force) measured by MD are always positive, and nonequilibrium processes are always observed in MD simulations to be dissipative, just as in the macroscopic real world.

In this paper, we will confront the doubtful status of the mathematics of nonlinear-response theory with the physical and mathematical reality of MD simulations of nonequilibrium steady states. Using formalism and notation[1] established for time-*dependent* response, we will derive an expression for the nonlinear response of an ensemble of nonequilibrium experiments, where the external driving force is time *independent.* Then we will derive the "Kawasaki expression"[4] for the nonequilibrium distribution function at the steady state and show that, mathematically speaking, this expression is ill defined and meaningless: new MD results show that the distribution collapses to a fractal object, the so-called "strange attractor,"[5] which cannot be expressed by a regular function. In spite of the singular (fractal) nature of the distribution function, the observable response *can* be expressed in terms of the transient (nonequilibrium) correlation function. In other words, even though the formalism is inadequate to describe the distribution, we are still justified in taking averages of mechanical properties as the steady state is approached and achieved. Before reviewing the mathematical highlights of classical response theory, let us first outline some important physical concepts.

In thinking about the mathematical description of the flow of trajectories in phase space, it is useful to imagine two kinds of propagators. Associated with the Lagrangian picture, or co-moving frame of reference that molecular dynamicists usually have in mind, is $U(t)$, the usual phase-space propagator, which moves an initial condition (denoted by Γ) to the phase point $\Gamma(t)$ along its trajectory: $\Gamma(t)=U(t)\Gamma$. Associated with the Eulerian picture, or space-fixed frame of reference, is $U^\dagger(t)$, the distribution-function "propagator," which is less intuitive, because it describes the evolution of f: $f(\Gamma,t)=U^\dagger(t)f_0(\Gamma)$. As we will show later, $U^\dagger$ acts like a "detective," tracing trajectories *backwards* in time to their origins.

These subtleties in U and $U^\dagger$ become truly significant in light of the fact that, even at equilibrium, physically interesting systems (many body, multidimensional, anharmonic—and therefore ergodic) exhibit Lyapunov instability, which means that nearby trajectories in phase space quickly (exponentially) diverge with time, typically within a few mean collision times. We illustrate in Fig. 1 the contrast between Lyapunov stability and instability for a blob of trajectories starting out in a phase-spac neighborhood. Lyapunov-stable flow can be represented by a continuous deformation—a "transformation" or gentle "mapping." On the other hand, Lyapunov-*unstable* flow (the interesting case in statistical mechanics) makes the blob quickly and grotesquely distort, fragment, and finally lose all sense of connectivity. Thus $U^\dagger$ must be a particularly clever detective to untangle the complex history of an ensemble of trajectories. The "transformation" or "map" is far from gentle, if such terms are even appropriate beyond the purely mathematical sense.

Keeping in mind the complexity that Lyapunov instability imposes on $U^\dagger$, we can nevertheless formally describe the time evolution of the nonequilibrium distribution function. The steady state is of particular interest, since the formalism simplifies. As we show later in the paper, the steady-field distribution function collapses from the initial regular, smooth equilibrium distribution

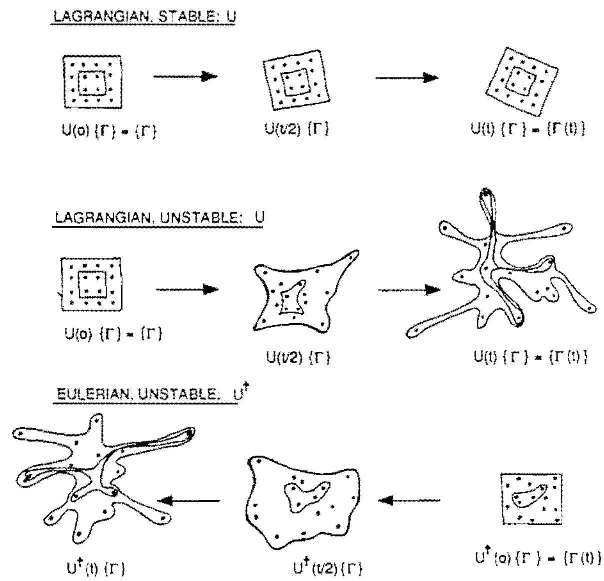

FIG. 1. Schematic representation of Lyapunov-stable and Lyapunov-unstable flow of a neighborhood of trajectories $\{\Gamma\}$ in phase space at times 0, $t/2$, and t. The top sequence shows, in the Lagrangian picture, a blob of points that distorts in such a way as to maintain local neighborhoods under Lyapunov-stable flow. The middle sequence, again in the Lagrangian picture, shows the fragmentation of a neighborhood of phase-space points under Lyapunov-unstable flow, the usual case in statistical mechanics; the action of the phase-space propagator U is to drive each trajectory forward in time (the fragmentation depicted here occurs even at equilibrium). The final sequence, in the Eulerian picture, shows a differential Eulerian box $d\Gamma$ centered at phase-point Γ (at the far right-hand side), where a series of trajectories $\{\Gamma\}$ have arrived at time t; the distribution-functio "propagator" $U^\dagger$ (like the detective in a mystery story) searches backwards in time to find the whereabouts of the trajectories (like suspects) at the time of the crime ($t=0$).

f_0 onto an object of lower, fractional dimensionality in the phase space—onto a fractal strange attractor. In order to visualize this collapse of the smooth initial distribution function, we subject the regular cloud of equilibrium points, representing an ensemble of MD experiments, to a steady external field, beginning at time zero. Inexorably, the face of the strange attractor emerges out of the featureless equilibrium cloud.

A useful cartoon of this condensation to a lower-dimensional attractor is provided by subjecting the simple harmonic oscillator to the dissipative Rayleigh–van der Pol external field. The phase space for this cartoon is two dimensional: q is the coordinate and p is the momentum. At time zero, the field ϵ is turned on, and the equations of motion are

$$\dot{q}=p \; ,$$

$$\dot{p}=-q-\epsilon(q^2-1)p \; .$$

In contrast to nonequilibrium statistical mechanics,[5] the Rayleigh–van der Pol equations of motion are intrinsically *time irreversible*. Nevertheless, it is instructive to see that the equilibrium (Gaussian) distribution for the harmonic oscillator, seen in Fig. 2 as a two-dimensional cloud, collapses under the Rayleigh–van der Pol dynamics onto a one-dimensional attractor, an odd-shaped racetrack. For a box near the attractor [see case (a) of Fig. 3], the number of trajectories rises exponentially, at least for a while. For a box far from the attractor, such as the origin [see case (b) of Fig. 3], the number of trajectories as a function of time drops qualitatively exponentially.

This behavior of the distribution function for a *non*dissipative statistical-mechanical system can be obtained from nonlinear-response theory in the form of the Kawasaki expression. As in the case of a dissipative system (such as Rayleigh–van der Pol), the attractor has lower dimensionality than the full space, and occupies zero volume. However, a statistical-mechanical system is different from dissipative ones characterized by limit cycles or fixed points (such as Rayleigh–van der Pol), in that the dimensionality of the attractor is noninteger, that is, the attractor is fractal like a sponge—full of holes.

The central message of this paper is that, because of the fractal nature of the nonequilibrium distribution function, the usual mathematical formalism has come to a dead end, at least as far as elucidating the properties of the distribution any further. However, the ensemble averages of the nonequilibrium response are well posed, because the trajectories themselves are perfectly well-defined objects, even though their distribution becomes a peculiar fractal object.

II. REVIEW OF CLASSICAL RESPONSE THEORY

The ensemble average of an observable function B of the phase space Γ at some particular time $t>0$ can be ex-

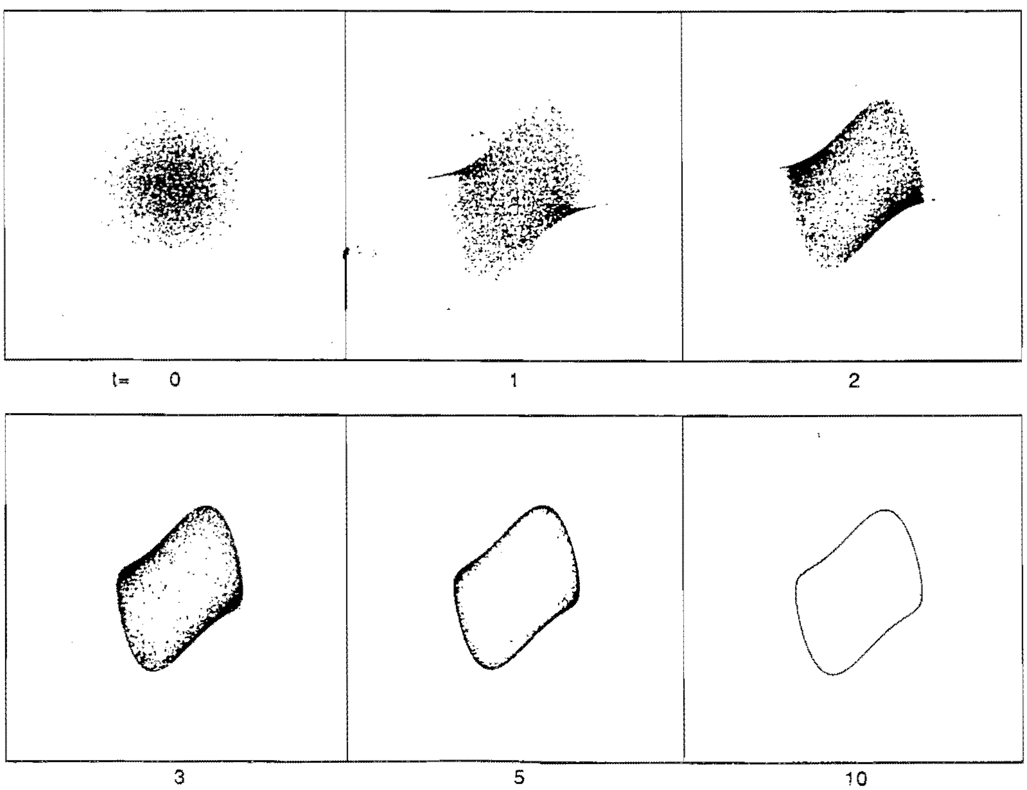

FIG. 2. Time evolution of 10 000 dissipative Rayleigh–van der Pol oscillators ($\epsilon=1$). The initial conditions are selected from a Gaussian of unit temperature; the mass and force constant of the unperturbed ($\epsilon=0$) one-dimensional harmonic oscillator are chosen to be unity, so that the period of one oscillation is 2π. The abscissa is the coordinate q and the ordinate is the momentum p; a 12×12 square centered at the origin is shown for the six times 0, 1, 2, 3, 5, and 10.

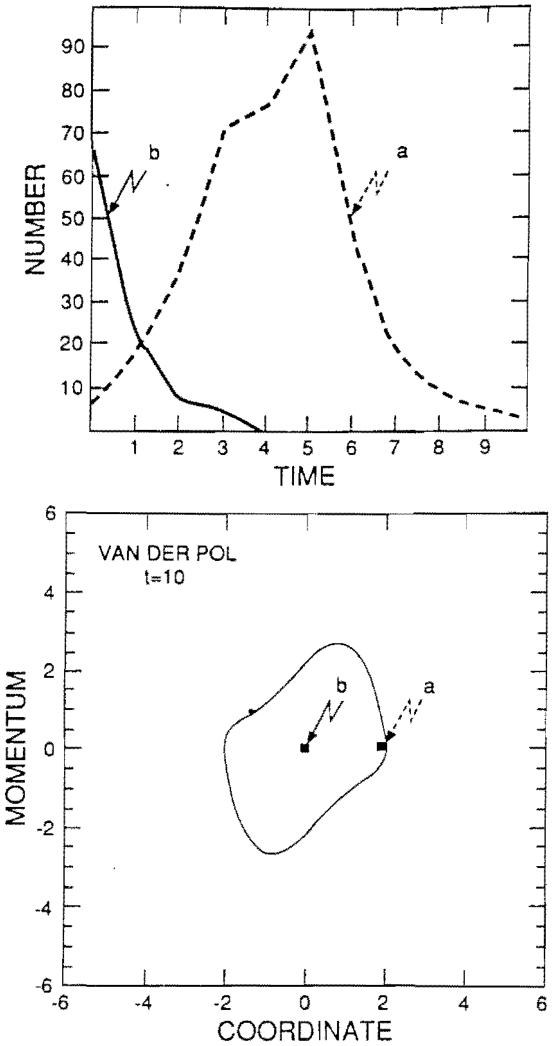

FIG. 3. Number of Rayleigh–van der Pol oscillator trajectories (out of a total of 10 000) that fall within two small, square boxes of side length 0.2, one near the attractor (marked *a*) and one near the origin (marked *b*), as a function of time. (The boxes are located in the $t=10$ snapshot, reproduced below the curves.)

pressed in the Eulerian (Schrödinger) picture, that is, in the space-fixed frame of reference. In the space-fixed frame, we position ourselves at a box $d\Gamma$ centered at a point Γ in phase space. (To establish the mathematical equivalence of the Eulerian picture and the Lagrangian, to be presented later, it is convenient to imagine that the initial conditions Γ are chosen at random from the equilibrium distribution f_0, with Eulerian boxes $d\Gamma$ constructed about each point Γ by erecting perpendicular bisecting planes between it and its nearest neighbors.) The observable evaluated at Γ remains constant in time, while the distribution function changes as ensemble elements pass through the differential box $d\Gamma$. We then sum over all boxes in phase space to obtain the ensemble average,

$$\langle B(t)\rangle = \int d\Gamma\, B(\Gamma) f(\Gamma,t) = \int d\Gamma\, B(\Gamma) U^{\dagger}(t) f_0(\Gamma) \; . \tag{1}$$

The equation of motion for the distribution-function "propagator" $U^{\dagger}$ is obtained from the Liouville equation (the phase-space continuity equation for the distribution function),

$$(\partial/\partial t) f(\Gamma,t) + (\partial/\partial\Gamma)\cdot[f(\Gamma,t)\dot{\Gamma}(\Gamma,t)] = 0 \; , \tag{2}$$

where the equations of motion for any trajectory inside the differential box $d\Gamma$ centered at the phase point Γ are denoted by $\dot{\Gamma}(\Gamma,t)$. The equations of motion presented in the example below are for mass flow under an external, time-varying field $X(t)$—for example, an electric field acting on charged particles, giving rise to the electrical conductivity. (Similar kinds of equations of motion are employed in MD simulations of a variety of hydrodynamic flows of momentum and energy, such as Couette shear flow, for example.[1]) The phase-space point is represented by $\Gamma=(q,p,\zeta)$, where q are the particle coordinates and p are the momenta. In order to maintain the temperature at a constant value T in the face of the driving force, we introduce a deterministic feedback mechanism, represented in the following by the Nosé-Hoover thermostat;[6,7] ζ is an additional dynamical thermostating variable (a dimensionless "friction" coefficient):

$$\dot{\Gamma}(\Gamma,t) = \begin{cases} \dot{q} = p/m & \text{(3a)} \\ \dot{p} = F(q) + X(t) - \nu\zeta p & \text{(3b)} \\ \dot{\zeta} = (\nu/g)\sum (p^2/mk_BT - 1) \; , & \text{(3c)} \end{cases}$$

where ν is a parameter that fixes the rate of thermostating, and the sum is over the g momentum degrees of freedom (k_B is Boltzmann's constant). The kinetic energy is $K(p)=\sum p^2/2m$; its ensemble average (as well as its time average along each trajectory, over times that are long compared to the response time of the thermostat ν^{-1}) is guaranteed by the thermostat to be $\langle K\rangle = \frac{1}{2} g k_B T$. The internal forces are obtained from the potential energy $\Phi(q)$: $F=-(\partial\Phi/\partial q)$. By setting both the external force and thermostating rate to zero, Hamilton's (Newton's) equilibrium equations of motion are recovered.

We emphasize that the particle equations of motion in externally driven systems need not be derivable from a Hamiltonian. Moreover, in order to achieve a steady state, deterministic thermostating (which also need not be derivable from a Hamiltonian) *must* be included somehow, either in boundary regions[5] or homogeneously throughout the sample, so as to prevent dissipative heating of the system.

Under Nosé-Hoover equilibrium dynamics, the long-time average of an observable is equivalent to a canonical-ensemble average.[7] This can be shown to be true for all nontrivial statistical-mechanical systems, since (1) such systems are ergodic and strongly mixing, and (2) the flow of trajectories under the equilibrium thermostated equations of motion $\dot{\Gamma}_0(\Gamma)$ [Eqs. (3), with $X\equiv 0$] satisfies the Liouville equation [Eq. (2)]. Thus any trajectory will eventually visit every box $d\Gamma$ located at Γ in phase space, with the probability $f_0(\Gamma)d\Gamma$ given by the equilibrium canonical distribution function,

$$f_0(\Gamma)=\frac{1}{Z}\exp[-\beta H_0(q,p)-\tfrac{1}{2}\varphi\zeta^2]\ ,$$

where the internal energy is $H_0=K+\Phi$, $\beta=1/k_BT$, and the partition function is $Z=\int d\Gamma\exp(-\beta H_0-\frac{1}{2}\varphi\zeta^2)$.

With $f(\Gamma,t)=U^\dagger(t)f_0(\Gamma)$, the Liouville equation [Eq. (2)] can be rewritten as

$$\begin{aligned}(\partial/\partial t)U^\dagger(t)f_0(\Gamma)&=-(\partial/\partial\Gamma)\cdot[\dot\Gamma(\Gamma,t)U^\dagger(t)f_0(\Gamma)]\\&=-iL^\dagger(t)U^\dagger(t)f_0(\Gamma)\ ,\end{aligned}\tag{4}$$

where the distribution-function Liouville operator $iL^\dagger$ operating on an arbitrary function B is defined by

$$\begin{aligned}iL^\dagger(t)B(\Gamma)&=(\partial/\partial\Gamma)\cdot[\dot\Gamma(\Gamma,t)B(\Gamma)]\\&=B(\Gamma)(\partial/\partial\Gamma)\cdot\dot\Gamma(\Gamma,t)\\&\quad+\dot\Gamma(\Gamma,t)\cdot(\partial/\partial\Gamma)B(\Gamma)\\&=B(\Gamma)\Lambda+iL(t)B(\Gamma)\ .\end{aligned}\tag{5}$$

The phase-space Liouville operator iL operating on an arbitrary function B is defined by

$$iL(t)B(\Gamma)=\dot\Gamma(\Gamma,t)\cdot(\partial/\partial\Gamma)B(\Gamma)\ ,\tag{6}$$

while the logarithmic expansion rate of phase-space volume is

$$\Lambda=(\partial/\partial\Gamma)\cdot\dot\Gamma(\Gamma,t)\ ,$$

We emphasize the fact that the phase-space expansion Λ in Eq. (5) is not an operator (with an attached i), but rather a multiplicative factor. For isolated Hamiltonian systems, $\Lambda\equiv0$ (the Liouville theorem); for thermostated (Nosé-Hoover) equilibrium systems, Λ fluctuates about zero; for nonequilibrium thermostated systems, $\Lambda=-\sum\nu\zeta$, where the sum is over all momentum degrees of freedom, with the possibility of independent thermostating variables ζ for boundary and bulk regions. The rate at which heat flows *into* the system[8] is $\dot Q=-2\nu\langle K\zeta\rangle$, which at the steady state is $k_BT\langle\Lambda\rangle$. Because we are interested in systems driven toward the nonequilibrium steady state, work is done on the system by external forces and heat extracted from the system by the thermostat. In such a case, Λ fluctuates about a negative constant at the steady state. Therefore, from the first law of thermodynamics, we are forced to conclude that, as a system is approaching the nonequilibrium steady state, the occupied phase space must *contract steadily* toward an object of lower dimensionality than the equilibrium distribution—the strange attractor.

When the external driving and thermostating are done in boundary regions separate from the bulk, rather than homogeneously throughout the sample, computer simulations strongly suggest[9] that, for small systems, the phase-space dimensionality loss induced by the boundary regions only rarely exceeds the extra dimensions required to describe them. Even though the dimensionality loss is associated with the thermostating variables ζ, the phase-space contraction cannot be localized in the thermostated boundaries. Rather, it is spread throughout the system, including the Newtonian bulk region, a point which is clarified by the example of heat flow: formally $\Lambda_N=0$ in the Newtonian region, $\Lambda_C<0$ in the cold boundary region, and $\Lambda_H>0$ in the hot region, but the total $\Lambda=\Lambda_H+\Lambda_N+\Lambda_C<0$. If the phase-space dimensionality were localized, then it would steadily increase in the hot boundary region—a clearly nonsensical result.

From Eq. (4), we see that the equation of motion for $U^\dagger$ is

$$(\partial/\partial t)U^\dagger(t)=-iL^\dagger(t)U^\dagger(t)\ ,\tag{7}$$

where the time derivative brings down $-iL^\dagger$ on the *left*-hand side of the "propagator." The solution to this equation of motion is therefore the usual (left-hand-sided) time exponential,

$$\begin{aligned}U^\dagger(t)&=\exp\left[-\int_0^t ds\, iL^\dagger(s)\right]\\&=\sum_{n=0}^{\infty}(-1)^n\int_0^t ds_1\cdots\\&\qquad\times\int_0^{s_{n-1}}ds_n iL^\dagger(s_1)\cdots iL^\dagger(s_n)\ .\end{aligned}\tag{8}$$

When one is thinking of a nonequilibrium molecular-dynamics trajectory, the ensemble average of an observable B is most naturally expressed in the Lagrangian (Heisenberg) picture, that is, in the co-moving frame of reference. In the co-moving frame, the observable along each trajectory changes with time, while the probability, or weight, of each trajectory remains fixed. The equivalence of the Eulerian and Lagrangian pictures can be verified by expanding $U^\dagger$ in Eq. (1), as shown in Eq. (8), and integrating by parts.[1] In the Lagrangian frame, $\langle B(t)\rangle$ is obtained by summing over all trajectories,

$$\begin{aligned}\langle B(t)\rangle&=\int d\Gamma\, f_0(\Gamma)B(\Gamma(t))\\&=\int d\Gamma\, f_0(\Gamma)U(t)B(\Gamma)\ .\end{aligned}$$

The phase-space propagator U operates on all occurrences of an initial phase point Γ, moving it along a trajectory to the new phase $\Gamma(t)$ at time t: $\Gamma(t)=U(t)\Gamma$. The particle equations of motion for spatially homogeneous driving are obtained from Eqs. (3) by interpreting $\dot\Gamma(\Gamma,t)$ as being a function of the *initial* phase Γ and the external driving force evaluated at time t, and then applying the propagator $U(t)$,

$$\dot\Gamma(\Gamma(t),t)=U(t)\dot\Gamma(\Gamma,t)\ ,$$

so as to match up the phase $\Gamma(t)$ with the external force $X(t)$.[1] In this way, we make the propagator operate only on the initial phase, so that observables evaluated along a trajectory can be expressed as functions of the initial phase.

Thus the equation of motion for the propagator U is obtained from the equation of motion of an observable B, assuming that B has no intrinsic time dependence ($\partial B/\partial t=0$),

$$(d/dt)B(\Gamma(t))=\dot\Gamma(\Gamma(t),t)\cdot[\partial/\partial\Gamma(t)]B(\Gamma(t))\ .$$

With $B(\Gamma(t))=U(t)B(\Gamma)$, this can be rewritten as

$$(\partial/\partial t)U(t)B(\Gamma)=U(t)\dot{\Gamma}(\Gamma,t)\cdot(\partial/\partial\Gamma)B(\Gamma)$$
$$=U(t)iL(t)B(\Gamma)\ ,$$

where the phase-space Liouville operator has already been defined in Eq. (6). Hence the equation of motion for U is

$$(\partial/\partial t)U(t)=U(t)iL(t)\ ,$$

which emphasizes that the time derivative brings down iL on the *right-hand* side of the propagator. The solution to this equation of motion is therefore a "right-hand-sided time-ordered exponential,"[1]

$$U(t)=\exp_R\int_0^t ds\, iL(s)$$
$$=\sum_{n=0}^{\infty}\int_0^t ds_1\cdots\int_0^{s_n-1}ds_n iL(s_n)\cdots iL(s_1)\ .$$

For times greater than that characteristic of the Lyapunov instability of trajectories in phase space, namely, $\lambda^{-1}\approx\tau_{\text{collision}}$ (typically on the order of the mean particle-particle or phonon-phonon collision time), the action of the distribution-function "propagator" $U^\dagger$ is quite different from that of the phase-space propagator U. Even at equilibrium, a blob of trajectories, initially clustered within a small, finite volume $\Delta\Gamma$ in phase space, maintains its identity for only a small, finite length of time (such that $e^{\lambda t}\Delta\Gamma$ is a constant) before Lyapunov instability fragments it irrevocably, as shown in Fig. 1. Therefore $U^\dagger$ traces back exponentially diverging trajectories that have reached the fixed Eulerian box at time t and gathers up the initial values of their trajectory weights. In this way, $U^\dagger$ acts more like a detective than the propagator U, in that it traces the footsteps of trajectories backwards in time.

In the Eulerian picture, using the equation of motion for $U^\dagger$ [Eq. (7)], we can rewrite Eq. (1) for the ensemble average of an observable,

$$\langle B(t)\rangle=\int d\Gamma\, B(\Gamma)\left[f(\Gamma,0)+\int_0^t ds\frac{\partial}{\partial s}f(\Gamma,s)\right]$$
$$=\langle B(0)\rangle-\int_0^t ds\int d\Gamma\, B(\Gamma)$$
$$\times iL^\dagger(s)U^\dagger(s)f_0(\Gamma)\ . \quad (9)$$

III. STEADY-FIELD DISTRIBUTION FUNCTION: KAWASAKI EXPRESSION

Now we introduce the simplifying features of the step-function external field, $X(t)=X\Theta(t)$. In this case, for $t>0$, $iL^\dagger$ becomes time independent and $U^\dagger=\exp(-itL^\dagger)$; consequently, for external fields that do not vary in time, the phase-space Liouville operator and propagator commute,

$$iL^\dagger U^\dagger(t)=iL^\dagger[1-itL^\dagger+(1/2!)(itL^\dagger)^2-+\cdots]$$
$$=[1-itL^\dagger+(1/2!)(itL^\dagger)^2-+\cdots]iL^\dagger$$
$$=U^\dagger(t)iL^\dagger\ .$$

By virtue of this commutativity for steady fields, we are able to rewrite the expression in Eq. (9) for the ensemble average of an observable:

$$\langle B(t)\rangle=\langle B(0)\rangle-\int_0^t ds\int d\Gamma\, B(\Gamma)U^\dagger(s)iL^\dagger f_0(\Gamma)$$
$$=\langle B(0)\rangle+\beta X\int_0^t ds\int d\Gamma\, B(\Gamma)$$
$$\times U^\dagger(s)J(\Gamma)f_0(\Gamma)\ , \quad (10)$$

where we have used the properties of the equilibrium canonical distribution function to obtain for the general time-dependent case[1]

$$iL^\dagger(t)f_0(\Gamma)=-\beta J(\Gamma)X(t)f_0(\Gamma)\ ;$$

here, $\dot{W}=-\langle J\rangle X$ is the rate of work done *by* the system on the external world,[3] which is balanced at the steady state by $\dot{Q}$, the heat flow *into* the system. The dissipative flux J is related phenomenologically to the driving force by the transport coefficient $\alpha>0$: $\langle J\rangle=\alpha X$, as in linear (Navier-Stokes) hydrodynamics. Nonlinear effects are ascribed to the transport coefficient itself, namely, $\alpha(X)=\alpha_0+$(higher-order terms in X). From the example equations of motion in Eqs. (3), which can be applied[10] to the case of a point particle of mass m diffusing through a Lorentz forest of infinitely massive scatterers under the driving force $X=mg$ (gravitational field in the x direction), the diffusive flux is the scattered particle's x velocity, $J=p_x/m=v_x$, which becomes αX at the steady state: α is the coefficient of mobility of the diffusing particle. For the example of Couette shear flow, the equations of motion are slightly more complicated than Eqs. (3) but similar in spirit.[1] The shear momentum flux is $J=-P_{xy}V$, which becomes $\eta V\dot{\epsilon}$ at the steady state: P_{xy} is the shear component of the pressure-volume tensor, V is the volume of the N-particle fluid, the external "field" is the imposed strain rate (the gradient of the fluid velocity), namely, $X=\dot{\epsilon}=(\partial/\partial x)u_y$, and the shear viscosity is η ($\alpha=\eta V$).

It is worthwhile to comment upon the physical meaning of Eq. (10). In the Eulerian picture, $U^\dagger$ gathers in the product Jf_0 from a multitude of trajectories. The result is then multiplied by the value of the observable in the Eulerian box and integrated over all such boxes. From Eq. (10), we see that the nonequilibrium distribution function can be expressed as

$$f(\Gamma,t)=f_0(\Gamma)+\beta X\int_0^t ds\, U^\dagger(s)J(\Gamma)f_0(\Gamma)$$
$$=f_0(\Gamma)+\beta X\int_0^t ds\, f(\Gamma,s)U^\dagger(s)J(\Gamma)$$
$$=f_0(\Gamma)\exp\left[\beta X\int_0^t ds\, U^\dagger(s)J(\Gamma)\right]\ , \quad (11)$$

where the last expression, derived by Yamada and Kawasaki over two decades ago[4] and referred to ever since as the so-called "Kawasaki expression," is the formal solution satisfying the Liouville equation [Eq. (2)], as can be verified by partial differentiation with time.

The physical interpretation of Eq. (11) is that the equilibrium Boltzmann factor in f_0 is modified by the energy dissipated in the nonequilibrium process, along the multitude of contributing trajectories. In this form of the Kawasaki expression for the distribution function, it is clear that, by analogy with its action upon the distribution function, the multitrajectory "propagator" $U^\dagger$, in its role as detective, searches *backwards* in time for original values of J.

It is by now well known from numerical simulations[5,10] that at a nonequilibrium steady state, the phase-space distribution function indeed collapses onto a "strange attractor" of lower dimensionality than the full equilibrium phase space (for examples, see figures of Pioncaré sections in Refs. 5 and 10). In light of this collapse, how are we to interpret Eq. (11)? Consider first an Eulerian box at $\Gamma=\Gamma^*$ near the strange attractor. (Figure 3 shows a cartoon example for the dissipative Rayleigh–van der Pol oscillator.) For $t \gg \lambda^{-1}$, that is for times greater than the Lyapunov time, all but an infinitesimal fraction of the ensemble of starting states (trajectories) will have collapsed onto the strange attractor. For most of the time interval from $s=0$ to t in the time integral in Eq. (11), the original values of J for the trajectories in the box at Γ^* will have come from points that lie in the vicinity of the attractor, while for times s in the range from $t-\lambda^{-1}$ to t they will have originated from almost anywhere, thereby contributing nothing of relative importance to the integral. The former (nonzero) contributions to the integral, however, are characteristic of the steady state, giving approximately $\langle J\rangle_{SS}t=\alpha Xt$. Therefore, at time t, for a box Γ^* "near" the attractor (that is, within a "distance" $d\Gamma$, in the sense that $e^{\lambda t}d\Gamma$ is a constant), the probability is qualitatively enhanced exponentially with time, that is, by $\exp(+\beta\alpha X^2 t)$. [See case (a) of Fig. 3.] This early-time exponential growth finally must give way to exponential decay, since the attractor (like a sponge) occupies zero volume: the inexorable collapse of all trajectories toward the attractor must continue forever.

What about those Eulerian boxes clearly *not* in the vicinity of the attractor? By symmetry, we can say that trajectories which have recently arrived in such a box must have come from very near the "strange *repellor*." The repellor can be constructed by taking the set of final limiting states that make up the attractor and applying the time-reversal transformation to it: coordinates $q\to q$ and momenta $p\to -p$. On the repellor, transport coefficients are negative and the second law of thermodynamics is violated, as pointed out by Loschmidt in his famous paradox.[5] (The resolution to this paradox of dissipative macroscopic behavior arising from atomistic time-reversible equations of motion, is that long-lived states with negative transport coefficients are unobservable in an ensemble of nonequilibrium experiments. Such states can arise only from initial conditions that begin on the repellor. But the probability of choosing such an initial condition at random is precisely zero, since the repellor occupies precisely zero volume in phase space.) Consequently, for the boxes that are almost everywhere in phase space *except* the measure-zero attractor and repellor, the probability drops, at least qualitatively, exponentially with time, that is, by $\exp(-\beta\alpha X^2 t)$. [See case (b) of Fig. 3, as well as the long-time behavior for case (a).]

It should by now be abundantly clear that the Eulerian picture, based upon the time evolution of the distribution function f, is without any practical utility (beyond formal manipulation) for nonequilibrium statistical mechanics. It is rendered useless in the approach to the steady state, where f becomes singular as it collapses onto the zero-measure strange attractor. This explains, for example, the Gibbs paradox, the peculiar divergence of the nonequilibrium entropy,

$$S(t)=-k_B\int d\Gamma\, f(\Gamma,t)\ln f(\Gamma,t)\ ,$$

which continues to drop toward minus infinity at the steady state.[2,8] Clearly, any functional of f itself, such as the above expression for the entropy, is meaningless at the steady state.

We wish to point out that Eq. (11) should not be misinterpreted[11] to mean that f becomes infinite *everywhere* in phase space when the nonequilibrium steady state is reached. Even though J approaches $\langle J\rangle_{SS}$ for all *trajectories* (except for the set of measure zero that start out on the repellor), f becomes singular only on the measure-zero attractor, in such a way that *the total probability is nevertheless conserved and equal to unity.* We see therefore that ad hoc "renormalization"[11] of f is both unjustified and unnecessary, and that any derivations which incorporate this erroneous view of the nature of the distribution must be viewed with great skepticism.

IV. STEADY-FIELD OBSERVABLES: TRANSIENT CORRELATION FUNCTION EXPRESSION

For observables, in contrast to direct properties of the distribution function itself, a computationally useful expression can be obtained from Eq. (10) by transforming from the Eulerian to the Lagrangian frame,[1]

$$\begin{aligned}\langle B(t)\rangle &= \langle B(0)\rangle+\beta X\int_0^t ds\int d\Gamma\, B(\Gamma)U^\dagger(s)J(\Gamma)f_0(\Gamma)\\ &= \langle B(0)\rangle+\beta X\int_0^t ds\int d\Gamma\, f_0(\Gamma)J(\Gamma)U(s)B(\Gamma)\\ &= \langle B(0)\rangle+\beta X\int_0^t ds\langle J(0)B(s)\rangle\ . \qquad (12)\end{aligned}$$

This is the so-called transient correlation function expression for the nonequilibrium response to a steady field, first derived by Visscher.[12] It cannot be too strongly emphasized, however, that Eq. (12) applies *only* for time-*independent* external fields; time-dependent fields destroy the commutativity of $iL^\dagger$ and $U^\dagger$ that allows simplification of Eq. (9). Note that in this derivation of Eq. (12), the temperature factor β arises naturally from Nosé-Hoover *dynamics*, rather than from the artificial construct of ensemble theory.

When we set the observable B equal to the dissipative flux J, whose equilibrium average is $\langle J(0)\rangle=0$, we obtain an expression from Eq. (12) for the field-strength dependent (nonlinear) transport coefficient, akin to the linear response Kubo formula,[1]

$$\alpha(X)=\frac{\langle J\rangle_{SS}}{X}=\frac{1}{k_BT}\int_0^\infty dt\langle J(0)J(t)\rangle\ . \qquad (13)$$

It should be emphasized that the transient correlation function $\langle J(0)J(t)\rangle$ is evaluated with the steady field turned on (as well as the thermostat); the *linear* response is obtained from the equilibrium correlation function

$\langle J(0)J_0(t)\rangle$ (here, the subscript 0 refers to the thermostated equilibrium dynamics, that is, with the field turned off, but with the thermostat on).

Equation (13) can be tested for a simple model of mass diffusion under gravitation: the two-particle Lorentz gas subject to periodic boundary conditions and thermostated at constant kinetic energy.[10] Since the flux J is the particle velocity v_x, the coefficient of mobility is $\alpha=D/k_BT$, where D is the nonlinear diffusion coefficient. At low densities, this problem can be solved exactly for two hard spheres in three dimensions (3D), and for two hard disks in 2D, by applying the Boltzmann equation. The agreement between the measured flux divided by the field and the time integral of the velocity autocorrelation function is good to at least three significant figures for a wide range of applied fields.[13]

Equation (13) can also be tested for the more realistic case of shear flow in dense fluids. When the external field (the strain rate) is weak, the test is made more difficult by the small signal-to-noise ratio. This problem can be overcome by applying the differential trajectory method of Ciccotti and co-workers:[14] along an equilibrium, Nosé-Hoover-thermostated trajectory, time origins are chosen at well separated (nonserially correlated) times, whereupon new, nonequilibrium, Nosé-Hoover-thermostated trajectories are generated by turning on the external field. The difference between the pair of simultaneous equilibrium and nonequilibrium trajectory segments is monitored at fixed response times beginning from the time origin, and the response is averaged over many such time origins. In work closely related to the differential-trajectory method, Morriss and Evans[15] have tested the differential trajectory form of Eq. (13) in the nonlinear regime and found that computing the differential transient correlation function is a useful procedure for very weak applied fields. In fact, they have been able to confirm that the shear viscosity η at very small shear rates is analytic in $\dot{\epsilon}$: $\eta(\dot{\epsilon})=\eta(0)-\eta_2\dot{\epsilon}^2+\cdots$, rather than nonanalytic, as Evans had previously conjectured.[16]

These results demonstrate that, at least for small, steady fields, the transient correlation function is especially useful in conjunction with differential trajectory methods. In summary, we see that, while the ensemble theory of Gibbs is of no particular use in nonequilibrium molecular dynamics, the Lagrangian picture of response theory, based on phase-space trajectories, maintains both its validity and utility.

V. CONCLUSIONS

We have presented a careful exposition of nonlinear response theory as applied to time-independent fields, demonstrating that the Kawasaki expression for the steady-state distribution function loses any meaning, because of the fractal nature of the distribution. We have also shown how the transient correlation function, naturally expressed in the co-moving (Lagrangian) frame, can yield useful results for the response, even though the distribution of trajectories becomes a fractal object. By now it should be quite clear, however, that the usual formalism, as represented by the Liouville continuity equation, cannot be applied blindly to fractal nonequilibrium distribution functions.

ACKNOWLEDGMENTS

This work profited greatly from discussions with Bill Ashurst, Berni Alder, Eddie Cohen, Denis Evans, Karl Kratky, Tony Ladd, and Gary Morriss. We also acknowledge discussions with Jerry Erpenbeck, Joel Lebowitz, and Bill Wood.

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

[2]J. W. Gibbs, in *The Collected Works of J. Willard Gibbs* (Longmans, New York, 1931), Vol. 2.

[3]See, for example, contributions in *Molecular-Dynamics Simulation of Statistical-Mechanical Systems*, edited by G. Ciccotti and W. G. Hoover (North-Holland, Amsterdam, 1986).

[4]T. Yamada and K. Kawasaki, Prog. Theor. Phys. **38**, 1031 (1967).

[5]B. L. Holian, W. G. Hoover, and H. A. Posch, Phys. Rev. Lett. **59**, 10 (1987); W. G. Hoover, H. A. Posch, B. L. Holian, M. J. Gillian, M. Mareschal, and C. Massobrio, Mol. Simul. **1**, 79 (1987).

[6]S. Nosé, Mol. Phys. **52**, 255 (1984); J. Chem. Phys. **81**, 511 (1984).

[7]W. G. Hoover, Phys. Rev. A **31**, 1695 (1985); B. L. Holian, in Ref. 3, p. 241; S. Nosé, Mol. Phys. **57**, 187 (1986).

[8]B. L. Holian, Phys. Rev. A **34**, 4238 (1986).

[9]W. G. Hoover and H. A. Posch, Phys. Rev. A **38**, 473 (1988), and unpublished results. Boundary-driven MD simulations were pioneered by Ashurst and summarized in an excellent review article (where the relationship of boundary-driven and homogeneously driven MD was also first discussed): W. G. Hoover and W. T. Ashurst, in *Theoretical Chemistry, Vol. I: Advances and Perspectives* (Academic, New York, 1975), 1.

[10]W. G. Hoover, J. Stat. Phys. **42**, 587 (1986); B. Moran and W. G. Hoover, J. Stat. Phys. **48**, 709 (1987); W. G. Hoover, B. Moran, C. G. Hoover, and W. J. Evans, Phys. Lett. A **133**, 114 (1988).

[11]D. J. Evans and G. P. Morriss, Mol. Phys. **61**, 1151 (1987); **64**, 521 (1988); G. P. Morriss and D. J. Evans, Phys. Rev. A **37**, 3605 (1988).

[12]W. M. Visscher, Phys. Rev. A **10**, 2461 (1974).

[13]W. G. Hoover, unpublished results.

[14]G. Ciccotti and G. Jacucci, Phys. Rev. Lett. **35**, 789 (1975); G. Ciccotti, G. Jacucci, and I. R. McDonald, J. Stat. Phys. **21**, 1 (1979).

[15]G. P. Morriss and D. J. Evans, Phys. Rev. A **35**, 792 (1987); D. J. Evans and G. P. Morriss, *ibid*. **36**, 4119 (1987).

[16]D. J. Evans, Phys. Rev. A **23**, 1988 (1981).