Resolution of Loschmidt's Paradox: The Origin of Irreversible Behavior in Reversible Atomistic Dynamics

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

William G. Hoover
Department of Applied Science, University of California at Davis, Livermore, California 94550, and
Department of Physics, Lawrence Livermore National Laboratory, University of California, Livermore, California 94550

and

Harald A. Posch
Institute for Experimental Physics, University of Vienna, A-1090 Vienna, Austria
(Received 4 May 1987)

We show that Nosé mechanics provides a link between computer simulations of nonequilibrium processes and real-world experiments. Reversible Nosé equations of motion, when used to constrain nonequilibrium boundary regions, generate stable dissipative behavior within an adjoining bulk sample governed by Newton's equations of motion. Thus, irreversible behavior consistent with the second law of thermodynamics arises from completely reversible microscopic motion. Loschmidt's reversibility paradox is surmounted by this Nosé-Newton system, because the steady-state nonequilibrium probability density in the many-body phase space is confined to a zero-volume attractor.

PACS numbers: 05.70.Ln, 05.20.−y, 46.10.+z, 65.50.+m

The resolution of Loschmidt's paradox, namely that particles obeying *reversible* equations of motion can (in accordance with the second law of thermodynamics) exhibit *irreversible* nonequilibrium behavior, is presented here in a novel application of Nosé mechanics.[1,2] This approach provides the formal structure needed to link nonequilibrium molecular-dynamics (NEMD) simulations of irreversible processes to real experiments. The picture we have in mind is a bulk sample composed of atoms governed by Newton's equations of motion; nonequilibrium boundary conditions and thermostatting are imposed in regions surrounding the bulk. In the boundary regions, the atoms are governed by reversible Nosé equations of motion, which, for example, might constrain the first and second moments of the velocity distribution so as to generate Couette shear flow at constant temperature. Another application would be heat flow, with the two boundary regions thermostatted at different temperatures. Figure 1 is a schematic of such a composite boundary-bulk-boundary system, with completely reversible dynamics achieved by Nosé-Newton-Nosé equations of motion, respectively. In this paper, we will show the relationship of this Nosé-Newton nonequilibrium system to the analogous boundary-driven NEMD introduced by Ashurst and Hoover[3] some fifteen years ago, as well as to the homogeneous NEMD equations of motion, which are non-Hamiltonian but nevertheless time reversible.[4] Finally, we point out the connection linking Nosé-Newton reversible mechanics with the (irreversible) second law of thermodynamics. We will emphasize nonequilibrium steady states in our presentation, but the generalization to cyclic processes involving overall dissipation is straightforward.

Nosé's recent modification of Hamiltonian mechanics[1,2,4–7] makes it possible to simulate the equilibrium dynamics of many-body systems with given values of the averaged temperature T or pressure P. The average can be carried out over time for a single system, or equivalently, for mixing systems, over an ensemble. Nosé showed that the long-time steady-state (equilibri-

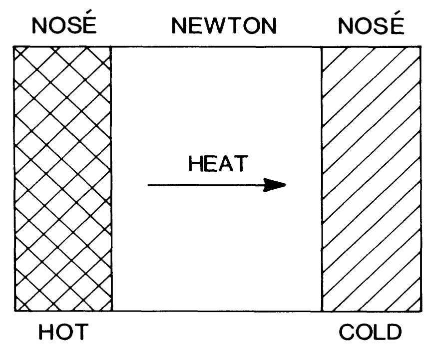

FIG. 1. Schematic of a boundary-bulk-boundary (Nosé-Newton) nonequilibrium system. Here, atoms in the left-most boundary region are Nosé thermostatted at a high temperature; atoms in the middle (bulk) region are governed by Newton's equations; atoms in the right-most boundary region are Nosé thermostatted at a low temperature. Heat flows through the bulk region from left to right. (Vertical walls separate particles in the three regions via elastic, specular collisions.)

um) distribution in phase space corresponds to the Gibbsian canonical or isothermal-isobaric ensemble in the constant-V and -T or constant-P and -T case, respectively. The distinguishing feature of Nosé's isothermal mechanics is that it allows us to maintain the temperature of particles in a boundary region at a value T through a *reversible* "friction" coefficient ζ. For these boundary particles, Nosé's deterministic reversible equations of motion relate the accelerations $\ddot{q}$ to the forces F, which depend on the coordinates q (p are the momenta):

$$\dot{p} = m\ddot{q} = F(q) - \zeta p. \tag{1}$$

The friction coefficient for a given boundary region satisfies the *integral* feedback equation

$$\zeta(t) = \zeta(0) + \int_0^t ds\, \dot{\zeta}(s), \tag{2a}$$

$$\dot{\zeta}(t) = [K(t)/K_0 - 1]/\tau^2. \tag{2b}$$

K_0 is the long-time steady-state average of the kinetic energy ($kT/2$ for each degree of freedom) in the given boundary region thermostatted at temperature T. τ is the response time of the thermostat, a parameter (in the limit that τ goes to infinity, Newton's equations of motion are recovered). The thermostatting, or friction coefficient ζ, fluctuates about zero at equilibrium, with fluctuations which vanish in the thermodynamic limit.

There is a close link between Nosé thermostatting and Gauss's principle of least constraint, a standard classical mechanics textbook method for implementing both holonomic and nonholonomic constraints.[4] If Gauss's principle is used to constrain the kinetic energy of a many-body system to a constant value K_0, exactly the same motion equations result [Eq. (1)]. But in this limiting case, with τ approaching zero, the reversible friction coefficient ζ_G is given explicitly:

$$\zeta_G = -\dot{\Phi}/2K_0, \tag{3}$$

where Φ is the potential energy. Thus, the Gauss isokinetic equations of motion are an example of *differential* feedback.

As pointed out in the introduction, the Nosé-Newton nonequilibrium system we have proposed provides the formal structure needed to link these computer simulations of irreversible processes to real experiments. Ashurst-Hoover boundary-driven NEMD is very closely related except that they used velocity scaling to thermostat the boundary regions.[3] (Velocity scaling approaches identically the reversible Gauss isokinetic equations of motion as the finite-difference time step is made smaller and smaller.) For situations close to equilibrium, Nosé-Newton mechanics reproduces the Green-Kubo results of linear-response theory.

Farther from equilibrium, Nosé thermostatting as well as external (non-Hamiltonian) forces can be used homogeneously throughout a system, so as to approximate hydrodynamic flows of mass,[8,9] momentum,[10] and energy,[11,12] in a way which is insensitive to system size. In these homogeneous systems we consider the boundary region to be our entire sample, with a single thermostatting coefficient applied to all degrees of freedom. (In the case of the relaxation of intramolecular vibrational modes,[12] the translational and rotational degrees of freedom are thermostatted at one temperature, while the vibrational modes are thermostatted at another; hence, *two* friction coefficients are required, one for each thermal reservoir.) Provided the flows are not *too* far from equilibrium,[10] these artificially constrained states are fairly accurate approximations to nonequilibrium steady states found in the absence of homogeneous constraints. Consequently, all these NEMD methods are related through the Nosé-Newton nonequilibrium system, demonstrating that irreversible dissipative behavior, consistent with the second law of thermodynamics, results from microscopic equations of motion which are completely time reversible in both bulk and boundary regions.

Exactly what is meant by "time reversibility?" This question has caused considerable confusion, not just in numerous informal discussions, but even in published works.[13] The fundamental test for, and definition of, time-reversible equations (which generate time-reversible motions) is that a *movie* of such a motion (that is, a record of the time dependence of the particle coordinates), run backwards through a movie projector, would still satisfy exactly the same equations of motion. The Nosé-Newton equations of motion *are* time reversible in this sense. Because in Nosé's original Hamiltonian derivation the friction coefficients ζ arise as momenta, all these friction coefficients, as well as all the particle momenta, change sign in the time-reversed motion. It is clear, in the typical equations of motion [Eqs. (1) and (2)], that changing the signs of the time t, the momenta p, and the thermostatting coefficients ζ, while leaving the coordinates q on which the forces F depend unchanged, generates the reverse trajectory.

This behavior is qualitatively different from that typical of chaotic dissipative maps, such as the Hénon map,[14] or from that characterizing the dissipative equations of continuum mechanics, such as the diffusion equation. In both these irreversible cases the equation of motion is clearly *invertible* (meaning that the past can be calculated from the present) but the *form* of the equation which describes the forward evolution is different from the form of the equation describing the backward evolution. The Hénon map, for instance, contracts phase-space area when iterated forward in time. The inverted map obtained from the Hénon map expands areas and so has a qualitatively different analytic form. Likewise, the solutions of the diffusion equation

$$\partial\rho/\partial t = D\nabla^2\rho \tag{4}$$

can be inverted (extrapolated backward in time), but only by changing the transport coefficient D (which is intrinsically positive) to a negative value.

The mathematical structures of dissipative maps and the hydrodynamic equations are inherently irreversible. The Nosé-Newton equations are different: They are time reversible. That is, the inverted equations which trace motion backward along any trajectory are identical to the equations describing the forward motion. The amazing thing about the Nosé-Newton equations applied to steady-state nonequilibrium systems is that they are found (numerically) to produce dissipative behavior, just as the invertible but time-irreversible dissipative maps of chaos theory and the irreversible partial differential equations of fluid dynamics do. This thermodynamically irreversible behavior occurs in both small and large systems. For any such system with any initial condition, Nosé-Newton mechanics leads to heat flow from high temperature to low temperature,[11] corresponding to a positive heat conductivity. Likewise, positive diffusion coefficients[8,9] and viscosity coefficients[10] result, even for systems involving only a few degrees of freedom.

The thermodynamic interpretation of this dissipative property is of general validity: The analog of Liouville's theorem, with use of Nosé-Newton mechanics,[7] becomes an equation for the time evolution of the phase-space density $f(q,p,\zeta,t)$:

$$\dot{f}=f\sum\zeta, \tag{5}$$

where the sum is over all boundary degrees of freedom. (In the example shown in Fig. 1, the sum contains many terms from many particles, but only two different classes of terms, and only two different friction coefficients, namely one for the hot boundary and one for the cold boundary.) The sum over friction coefficients (ζ) can then be related to

$$\dot{E}=T\dot{S}=-kT\zeta, \tag{6}$$

the rate at which energy E is being exchanged between a given degree of freedom, thermostatted at temperature T, and its corresponding Nosé reservoir. If a reservoir extracts heat, the corresponding $\dot{S}$ is negative, increasing f; if a reservoir furnishes heat, the corresponding $\dot{S}$ is positive, decreasing f.

Thus if energy is dissipated in the steady-state system, then the sum of $\dot{S}$ over all reservoirs is negative, and the distribution function must eventually diverge to infinity at the steady state, indicating a collapse of the phase-space probability onto a subspace with zero volume. In the cases which have so far been analyzed numerically, the collapse does occur; the resulting subspace is indeed a fractal[15] attractor. The simplest such example[9] is shown in Fig. 2. (To date, the number of degrees of freedom, or dimensionality of phase space, has necessarily been restricted to few-body problems. Lyapunov spectra for eight particles in a three-dimensional fluid, driven by an external field, confirm the one-body and two-body results[8,16] of phase-space contraction to a fractal strange attractor.[17])

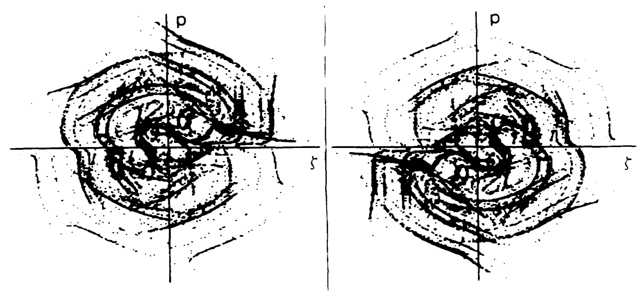

FIG. 2. Left: A Poincaré attractor section, corresponding to the relative probability of the momentum p and Nosé friction coefficient ζ, at a fixed value of q. There is a periodic sinusoidal potential and an external driving field. For details, see Ref. 9. Right: The corresponding repellor (with the momentum p and friction coefficient ζ changed in sign relative to the attractor) which, although violating the second law of thermodynamics, has such a small probability of being observed (precisely zero) as to be unobservable. The dimensionality of the Poincaré section is 1.50 ± 0.02. The prevalence of positive-p points implies that the conductivity is positive. The prevalence of positive-ζ points implies that the thermodynamic dissipation of work into heat is likewise positive.

How is this collapse of probability onto a zero-volume attractor related to the second law of thermodynamics? The phase-space states which can violate the second law by steadily converting heat back into work are precisely those of the corresponding unstable repellor (an object just like the stable attractor, but with the signs of the momenta, friction coefficients, and Lyapunov coefficients all changed). For the sinusoidal diffusion example, the attractor and repellor are illustrated on the left- and right-hand sides of Fig. 2. It is clear from Fig. 2 that the repellor states correspond both to an unphysical negative conductivity and to a negative dissipation, through which heat is continuously converted into work.

Thus steady states which *could* violate the second law, *if* they were observable, span a volume of exactly the same size as does the zero-volume attractor. By the choice of any state near the attractor which has been propagated forward in time for a time t_{forward}, then a change of the signs of the momenta and thermostatting coefficients, and propagation backward, the second law can be violated for a time t_{forward}. But the only way that a *permanent* (steady-state) violation of the second law could occur would require an inversion of a state precisely on the zero-volume attractor. These states occupy precisely zero volume and require an infinitely long simulation in the forward time direction for their characterization.

The conclusion of this novel analysis of Nosé-Newton mechanics is most interesting, and follows along the line of Prigogine's attempts to understand the irreversibility of the second law of thermodynamics through the structure of microscopic reversible equations.[18] (He views the problem of irreversibility from a different, but complementary, perspective, namely, the relaxation of a nonequilibrium initial condition toward equilibrium. The Nosé-Newton formalism applies here, too: We can imagine a nonequilibrium steady state having been achieved, and then turning *off* the driving force. The Newtonian bulk, whose distribution function is initially a zero-volume strange attractor, then relaxes toward equilibrium with its phase space expanding and the entropy increasing—irreversibility.) The reversibility paradox[6,19] disappears when Nosé-Newton mechanics is used to describe steady-state nonequilibrium systems, despite their mathematical reversibility. Any initial conditions which could violate the second law of thermodynamics have *precisely zero* probability, even for small systems with only a few degrees of freedom. Thus, the present combination of (1) the fractal concepts popularized by Mandelbrot, (2) the reversible dynamics introduced by Nosé (and related to Gauss's principle), and (3) computers powerful enough to study the consequences of these ideas, has resolved the old reversibility paradox for nonequilibrium steady states. That is, unstable states going backward in time are never observed, *not* because they violate the equations of motion, but rather because the probability of observing them is precisely zero.

It is a pleasure to acknowledge the help of many colleagues in coming to this understanding of the consistency of the second law of thermodynamics with reversible mechanics, in particular, Bill Ashurst, Giovanni Ciccotti, Denis Evans, Mike Gillan, Tony Ladd, Michel Mareschal, Carlo Massobrio, Bill Moran, Gary Morriss, and Rudy Pierleoni. Helpful suggestions from Willy Moss, Farid Abraham, and Francis Ree contributed significantly, too. One of us (B.L.H.) performed the work at Los Alamos under the auspices of the University of California, under U.S. Department of Energy Contract No. W-7405-Eng-36. Another of us (W.G.H.) performed the work at Livermore under the auspices of the University of California, under U.S. Department of Energy Contract No. W-7405-Eng-48.

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

[2]W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).

[3]W. T. Ashurst and W. G. Hoover, Phys. Rev. A **11**, 658 (1975).

[4]For a recent compilation of NEMD techniques, see W. G. Hoover, *Molecular Dynamics,* Lecture Notes in Physics Vol. 258, edited by H. Araki *et al.* (Springer-Verlag, Heidelberg, 1986).

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

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

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

[8]B. Moran, W. G. Hoover, and S. Bestiale, J. Stat. Phys. (to be published).

[9]W. G. Hoover, H. A. Posch, B. L. Holian, M. J. Gillan, M. Mareschal, C. Massobrio, and S. Bestiale, Molecular Simulations (to be published).

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

[11]D. J. Evans and W. G. Hoover, Annu. Rev. Fluid Mech. **18**, 243 (1986).

[12]B. L. Holian, J. Chem. Phys. **84**, 1762 (1986).

[13]D. C. Wallace, in *Shockwaves in Condensed Matter,* edited by Y. M. Gupta (Plenum, New York, 1986), pp. 41, 42.

[14]H. G. Schuster, *Deterministic Chaos* (Physik-Verlag, Weinheim, Federal Republic of Germany, 1984), pp. 96–98.

[15]G. P. Morriss, Phys. Lett. A (to be published).

[16]W. G. Hoover and H. A. Posch, Phys. Lett. A (to be published).

[17]B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).

[18]I. Prigogine, in *Fluctuations and Sensitivity in Nonequilibrium Systems,* edited by W. Horsthemke and D. K. Kondepudi (Springer-Verlag, Berlin, 1984).

[19]For a recent popular account, see T. Rotham, Discover **8**, No. 2, 63 (1987).