ELSEVIER

Physica D 112 (1998) 225–240

PHYSICA D

Time reversibility in nonequilibrium thermomechanics

Wm.G. Hoover [1]

Department of Applied Science, University of California at Davis-Livermore, Livermore, CA 94551-7808, USA

Abstract

"Time reversibility" of dynamical systems has had many different interpretations, for the governing differential equations, for the numerical approximations to their solutions, for the maps which caricature these equations and their solutions, and for the phase-space distributions describing ensembles of systems. Here I discuss some of these views, taken from recent research literature, and relate them to my own experience, which is primarily computational. I emphasize the usefulness of "thermomechanics", a simple generalization of conventional molecular dynamics, which leads – directly, inexorably, and paradoxically – from Hamilton's Principle of Least Action to the multifractal time-irreversible phase-space behavior summed up by the macroscopic Second Law of Thermodynamics.

Keywords: Time reversibility; Molecular dynamics; Fractals

1. Introduction

The fundamental differential equations of mathematical physics – Einstein's, Hamilton's, Lagrange's, Maxwell's, Newton's, and Schrödinger's – are "time reversible" [1,2]. In this they differ from thermodynamics. What does "time reversibility" mean? Simply that "solutions" of the fundamental equations – time histories of particle or field variables – can be followed either forward or backward in time without any change in the equations themselves. To "reverse time" only changes in the initial conditions are required.

The oldest of all these systems of time-reversible equations, those describing classical mechanics, are closest to our experience and are also the most easily understood. These days classical particle mechanics is changing, so as to incorporate thermodynamics and continuum mechanics within it. The old view of mechanics as formal and analytical is being replaced by a broader picture where singularities abound and "proofs" of new ideas are often computational. These changes are due to the rapidly expanding possibilities for computer simulation of complex systems, through the development of special algorithms and boundary conditions. The resulting methods and ideas are natural consequences of the computers which made them possible. I emphasize here generalized forms of the classical equations of atomistic point mechanics, selected for their usefulness in simulating chaotic nonequilibrium

[1] E-mail: hoover@bonampak.llnl.gov.

PII S0167-2789(97)00212-1

systems. I follow Feynman [3] in choosing Hamilton's Principle of Least Action as the conceptual basis for this approach.

Atomistic computer simulation has proceeded in stages. First, it was necessary to establish the agreement between dynamic and statistical averages, which is the basis of statistical mechanics, and to study the approach to equilibrium, which is the basis of kinetic theory [4]. Then, time correlation functions and linear transport coefficients were studied [5]. These early studies followed the motion of up to about one thousand atoms. In the course of this work, it became clear that any discrepancies between the various microscopic and macroscopic approaches could be reliably interpreted in terms of surface, or boundary, effects and fluctuations. Simulations of more complicated far-from-equilibrium flows, followed naturally. Typical examples include strong shockwaves [6], the brittle and ductile fracture of solids [7], and multiphase fluid flows [8]. Two recent efforts, involving millions of atoms, with special boundary conditions, are illustrated in Figs. 1 and 2.

About 10 years ago the understanding of nonequilibrium systems underwent a conceptual advance. New time-reversible motion equations [9,10] were developed. These "thermostatted" equations, incorporating "thermostat

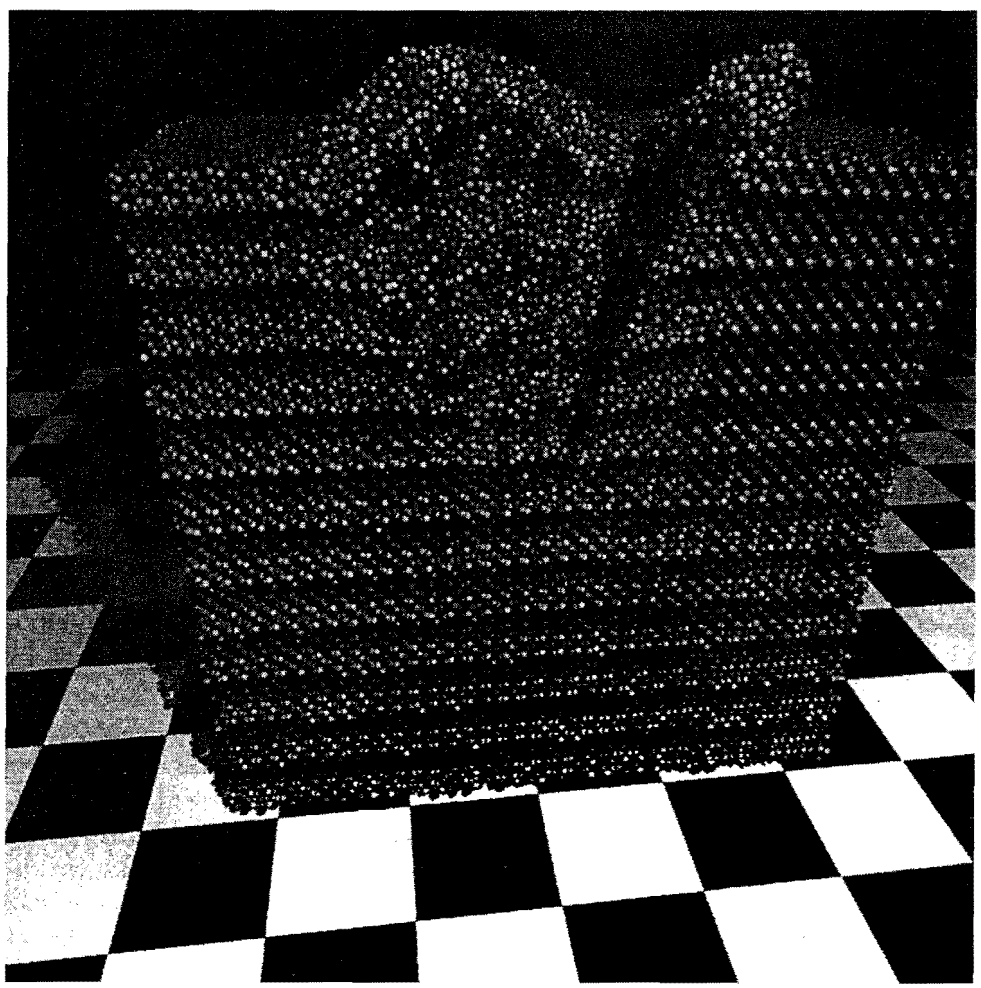

Fig. 1. Cutaway view of a simulated silicon crystal undergoing indentation, using a tetrahedral indentor. About one million atoms were used. See [39].

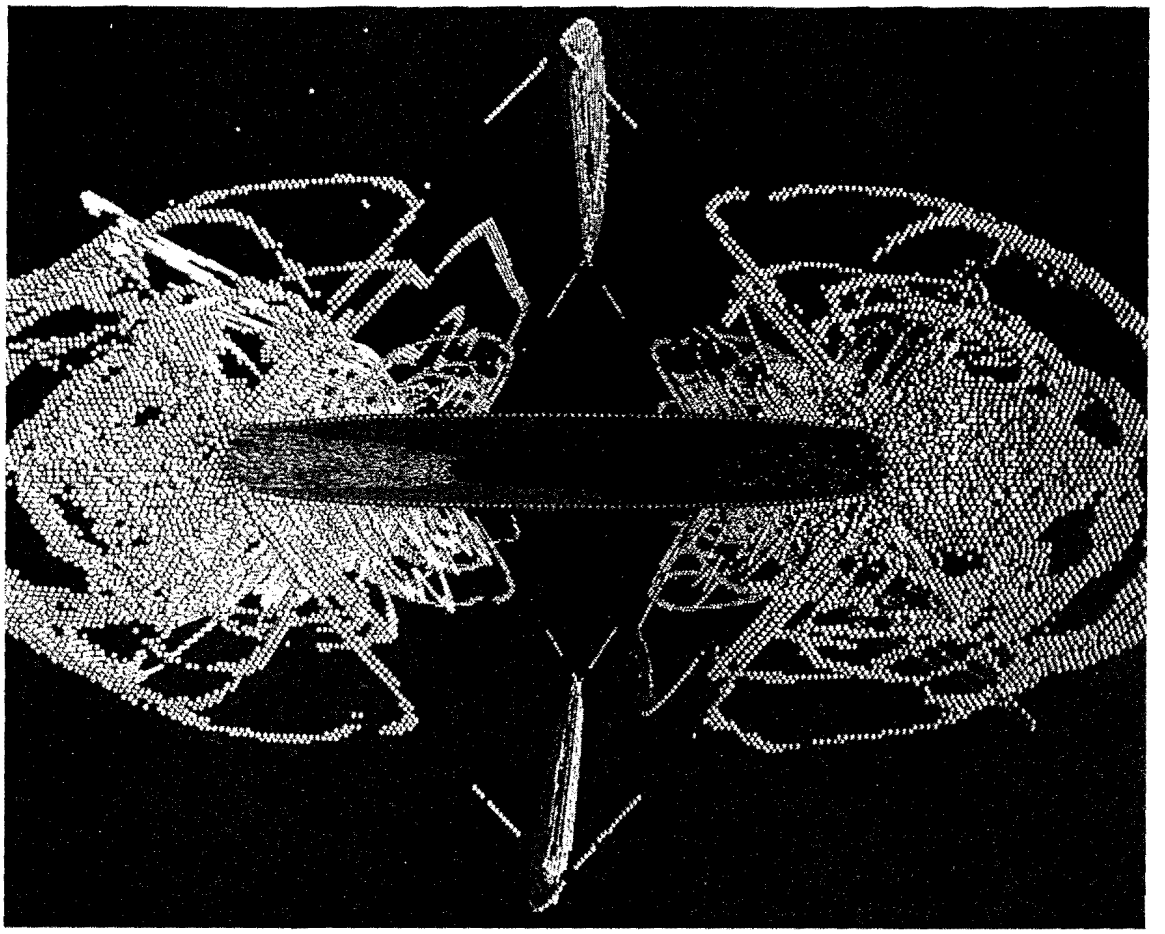

Fig. 2. Energetic copper atoms in the vicinity of an elliptical crack. The display locates crystalline dislocations emitted in the stress relaxation process. This simulation used 35 million atoms. See [7].

forces", describe the interaction of systems with momentum and energy reservoirs. In the laboratory, systems can be exposed to a variety of boundary, constraint, and driving forces. In simulations too, there is a wide variety of possible thermostat forces, constraining or driving the temperature, the energy, the pressure, or the heat flux of selected degrees of freedom. In Sections 3 and 4 I describe and discuss example thermostat forces based on Gauss' and Hamilton's variational principles. These variationally based forces are also the simplest, in the sense of Occam's Razor, that reproduce the exact results of Green–Kubo–Onsager linear-response theory. Farther from equilibrium, where linear-response theory fails, intercomparisons of the nonlinear responses show very little sensitivity to the detailed nature of the underlying thermostat forces [11].

The resulting nonequilibrium simulations [12–15] soon established that typical flows correspond to multifractal Lyapunov-unstable phase-space structures. These fractal structures are qualitatively different to the smooth distributions discovered by Boltzmann, Gibbs, and Maxwell. The nonequilibrium studies have shown that statistical ensembles of closed thermostatted nonequilibrium systems always generate heat, as a consequence of dissipation, and that extracting this heat results in a multifractal phase-space structure [14–16]. Thus, patently time-reversible motion equations, with thermostats, invariably generate irreversible behavior. This approach has forged strong links

between microscopic mechanics, macroscopic continuum mechanics, and thermodynamics. The thermostatted motion equations are firmly rooted in classical Hamiltonian mechanics [17,18].

There is a parallel effort, even more recent, based on the analysis of "open" nonequilibrium systems – systems which can exchange mass, as well as momentum and energy, with their environments – using either purely Hamiltonian mechanics [19–21], or hybrids including thermostats. Experience with equilibrium simulations has shown that the open-system approach is not very readily adapted to large-scale simulations. So far, the advocates of open systems have not faced up to the problem of constructing nonequilibrium particle sources useful outside the regime of linear transport theory. Nevertheless, the idealized studies so far carried out are of considerable pedagogical value in revealing features common to both open and closed systems, at, and away from, equilibrium. Both the approaches just mentioned are time reversible and lead to the fractal phase-space structures which indicate irreversible dissipation.

Here I review the main points of these developments in mechanics, emphasizing features which exhibit, simultaneously, interesting mathematical and physical features. I begin with a discussion of time reversibility from my perspective, computer simulation. I describe the way in which temperature, and thermostats, enter into mechanics, through time-reversible modifications of the usual equations of motion. I stress recently discovered links between the thermostatted motion equations and Hamiltonian mechanics, including the new emphasis on open systems. I use a few representative worked-out examples to illustrate these points. I then discuss the areas in which mathematicians and physicists could usefully collaborate, in grappling with the interpretation, correlation, and extension of the new ideas.

2. Time reversibility

Though time reversibility seems to be a familiar intuitive notion, the details can be confusing [2], and the definitions slippery, when it comes to applications. In part, this stems from the fact that dynamics is described by time-reversible differential equations and is imagined as the continuous time development of a set of precisely defined coordinates. Such continuity and precision are never both present in any interesting "solution" of the equations because such a solution is necessarily a finitely expressible "coarse-grained" numerical approximation. Such finite approximations often lack the time reversibility of the underlying differential equations. Approximate numerical solutions, when they attempt to describe functions of a continuous time variable, inevitably involve some approximation or truncation. Series of coordinate sets, given at discrete times, truncated Taylor or Fourier series in the time, and movies of the motion, are examples. It appears that the time reversibility of a particular approximate numerical solution depends not only on the form of the underlying differential equations, but also on the way in which the solution is approximated and presented. The picture becomes cloudier still when the evolution of a fractal-density ensemble of systems is described [22].

Let us focus on the simplest case, not an ensemble, but rather the evolution of a single dynamical system. Consider the following representative forms which dynamical "solutions" can take:

(i) a set of movie frames;
(ii) an ordered set of integer coordinates;
(iii) an ordered set of floating point coordinates;
(iv) a truncated Fourier time series;
(v) a truncated Taylor time series.

Evidently any movie is "time-reversible" in the simplest sense, whether or not the underlying equations or map (if any) are reversible, because the frames can be projected in reverse order. For differential equations this type of reversal is equivalent to changing the sign of the timestep $\Delta t \rightarrow -\Delta t$ in using a computer algorithm to approximate

the solution of a differential equation. This type of reversal has been called "invertibility" and "reversibility with time inversion" [1,2]. Such reversals are not at all what physicists mean by "time reversibility". For a physicist, the differential equations describing the forward evolution in time are necessarily identical to those going backward in time. Only the initial conditions, not the equations, differ in the reversed flow.

Let us consider the simplest possible finite description of a system's time evolution: integer coordinates given at equally spaced times [23]. Sets of such integer coordinates, describing the "time" development of a dynamical system, can be generated with a "bit-reversible" algorithm, such as the Størmer–Verlet centered-difference scheme:

$$\{q_+ - 2q_0 + q_- = (F/m)_0(\Delta t)^2\}.$$

Here the subscripts indicate three contiguous times, separated by intervals Δt. The right-hand side is to be truncated to an integer. This symmetric algorithm can evidently be extended, equally well, either forward or backward in time, and is patently time reversible. Any bounded solution must also be, in principle, periodic. Eventually, because the number of state points is finite, the initial state must repeat, so that the "dynamics" consists of a single periodic orbit.

If exactly this same algorithm were used with floating point, rather than integer, arithmetic, Lyapunov instability would typically destroy the exact step-by-step reversibility of any approximate trajectory, by exponential amplification of the inevitable roundoff errors. Whether or not this distinction between periodic reversible orbits and their irreversible relatives has any practical consequences is not known.

Although in practice simple stepwise algorithms are usual, in principle there are many alternative representations of dynamical trajectories. Truncated series expansions are among these possibilities. Truncated Fourier series suggest the absence of Lyapunov instability, because such series imply a periodic solution, while truncated Taylor series are obviously limited to trajectories of finite length. Neither of these series representations is specially useful for real problems, in which the time derivatives of the accelerations are often singular.

The most interesting, puzzling, and enduring aspect of time reversibility is the connection between atomistic microscopic dynamics and the macroscopic Second Law of Thermodynamics. Most explanations for this connection, though without doubt correct, are somewhat limited in scope. Boltzmann's dilute-gas approach, and the objections to it by Zermélo and Poincaré, are the most familiar [2,24]. The much more recent linear-response theory of transport of Green and Kubo is a first-order perturbation theory. Objections to Green–Kubo theory [25] are less convincing and not so well organized as those Boltzmann faced. Perhaps they are also less important, because the Green–Kubo–Onsager approach is limited to states which are close to equilibrium. It might well be thought that time-reversible equations of motion are inconsistent with the dissipative shrinking to the strange attractors that characterize nonequilibrium problems. The ways in which this can occur have become much clearer during the past 10 years and represent new paths to understanding the reversibility paradox faced earlier by Boltzmann, Green, and Kubo.

One single example [26], based on time-reversible maps [27], can illustrate the connection between reversible dynamics and irreversible behavior. The closest analog of a physicist's notion of time reversibility, for a two-dimensional map M, is to regard the "coordinate" x as unchanged by reversal, while the "momentum" y changes sign:

$$\{x, y\} \stackrel{M}{\to} \{x', y'\} \;\Rightarrow\; \{x', -y'\} \stackrel{M}{\to} \{x, -y\}.$$

Consider three such "time-reversible" maps of the unit square into itself,

$$\{-1/2 < x, y < +1/2\} \stackrel{X,Y,P}{\to} \{-1/2 < x', y' < +1/2\},$$

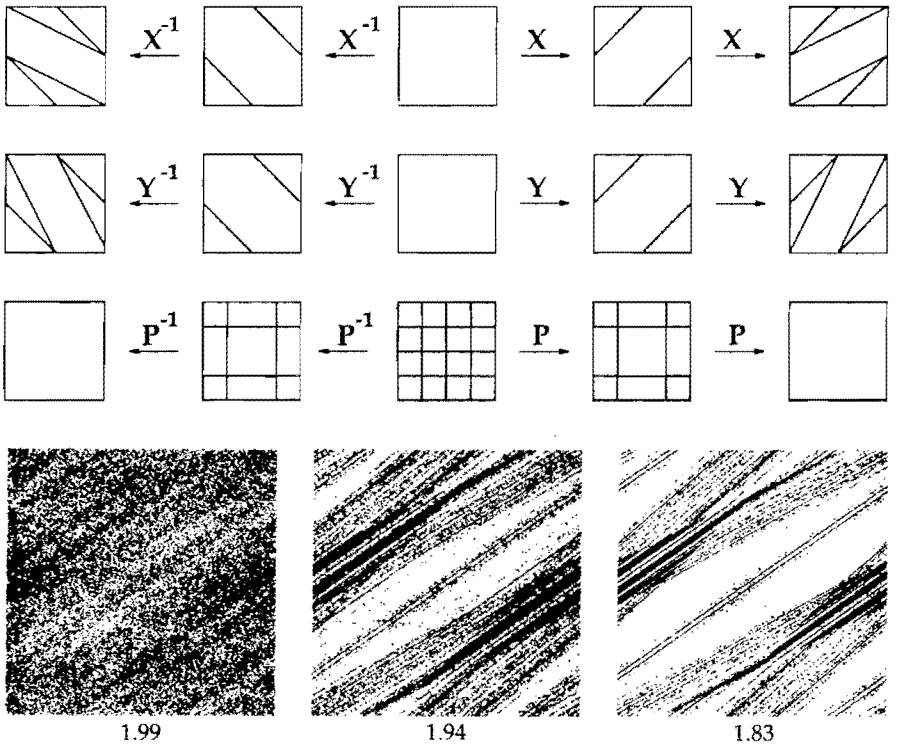

Fig. 3. The effect of three time-reversible mappings, a horizontal periodic shear X, a vertical periodic shear Y, and a twofold mirror reflection P, here shown for $m = 1/4$, are indicated. Attractors, shown with their fractal (information) dimensions, correspond to m values of 0.24, 0.20, and 1/6. See [26].

as illustrated in Fig. 3: (i) a horizontal periodic shear X, with the x displacement proportional to the y coordinate, while that y coordinate is left unchanged; (ii) a similar vertical periodic shear Y, in which the x coordinate is unchanged; and (iii) a twofold nonlinear reflection P, from both the closest horizontal mirror, at $x = \pm m$, and the closest vertical mirror, at $y = \pm m$. The horizontal and vertical shearing maps, the reflecting maps, and even any symmetric combination map, such as $XYPYX$, are "time-reversible" in the sense just described. The combination map $XYPYX$, when iterated, typically generates an ergodic Lyapunov-unstable multifractal distribution provided that m is not too far from 0.25.

Three examples [26], for three values of the mirror locations, $x = \pm m$; $y = \pm m$, are shown in Fig. 3. Each of these mappings links a "repellor" source to a similar "attractor" sink. The attractors, shown in the figure, are mirror images of their corresponding repellors. To visualize the corresponding repellors, simply "reverse time" – or, equivalently, change the sign of the momentum, y – by reflecting the attractors about the x-axis. The discontinuous nature of these piecewise-linear mappings, and the simplicity of their local Lyapunov exponents, could be made more "realistic" at the cost of additional nonlinear complexity. Nevertheless, these simple example maps already capture much the same paradoxical character as is shown by the ordinary differential equations typifying the time-reversible extensions of mechanics discussed in this paper. It takes only a few minutes' study to verify that the fractal and reversible natures of the maps are shared with those of a generalized rotated Baker Map, where the usual textbook mapping is rotated 45°.

Some of the conceptual difficulties associated with irreversibility have a semantic basis. For instance, thermodynamics contributes to the difficulty of our task by using the word "reversible" in a completely different sense.

The thermodynamic use describes relatively slow nonequilibrium processes, which are only infinitesimally far from equilibrium and which have negligible hysteresis. In thermodynamics an "irreversible" process is a process with noticeable dissipation.

Some of the difficulties are more basic. It is hard to imagine a computationally useful description of a fractal flow through phase space, particularly in the case that the initial distribution is fractal too. This apparent difficulty [22] can be avoided by considering coarse-grained distributions, reflecting the limited resolution, or information content, of any conceivable measurement or simulation [21].

In the remainder of this paper, I discuss links between the microscopic time-reversibility of atomistic dynamics and the macroscopic thermodynamic irreversibility described by the Second Law of Thermodynamics. From the computational standpoint, the simplest way to forge the link makes use of a mechanical temperature concept, with heat reservoirs [14,15]. Temperature [28], and thermostat forces, can be avoided, temporarily, by studying open systems [19–21]. However, the sources and sinks of particles, which drive these systems, must eventually be given operational significance, and that will require thermostats. For linear transport processes it can be confidently expected that the closed-system and open-system results will agree [21]. A still different approach seeks to describe dynamical averages through phase-space averages over sets of unstable periodic orbits [29]. This is by far the most cumbersome of the methods used to link reversible mechanics to macroscopic behavior.

In what follows I begin by describing the relation of temperature and heat transfer to the question of time reversibility. I then review some recently developed computational methods for implementing thermal constraints. Next, I focus on a peculiar consequence of these thermomechanical methods: the multifractal nature of the resulting phase-space distributions. Then I link the microscopic Principle of Least Action to the motion equations underlying the macroscopic Second Law. A few example problems and some remaining challenges for mathematicians and physicists conclude my contribution. Throughout, I concentrate on time reversibility as a means of coordinating these topics. I greatly admire Percy Bridgman's point of view, which stresses that interesting physical properties must have operational significance [30]. This requirement ensures consistency between the mathematical and physical points of view. Here, I make an effort to follow Bridgman's advice, using none but well-defined operational concepts.

3. Temperature, heat transfer and dynamic constraints

Equilibrium thermodynamics is based on the Zeroth-Law notion of temperature, and leads to a Second-Law state function, entropy, intimately related to temperature through the transfer of heat. The only useful conceptual basis for temperature is an operational one, suitable for use in nonequilibrium mechanical simulations, and based on the "ideal-gas thermometer", a collection of mass points so small, numerous, and frequently colliding that they are characterized by the equilibrium Maxwell–Boltzmann velocity distribution $\propto e^{-mv^2/2kT}$, where k is Boltzmann's constant, m the ideal-gas particle mass, and T is the corresponding temperature of this thermometer. Routine kinetic-theory calculations show that such idealized thermometers have two effects, a viscous drag force and a transfer of heat. On the average thermometers give up energy ("heat") *to* massive bodies whenever these bodies are cooler, and take it *from* them when they are hotter. On the average, a D-dimensional mass point, with $M \gg m$ and $V \ll v$, extracts heat *from* a reservoir at T whenever $MV^2 < DkT$, and releases heat *to* the reservoir, on the average, whenever $MV^2 > DkT$. This straightforward calculation suggests that a general definition of temperature be based on the mean squared velocity.

In nonequilibrium simulations it is not necessary to follow the details of the thermometric system-bath collisions. Heat reservoirs can be introduced into mechanics implicitly, rather than explicitly, by constraining the kinetic energy of selected "thermostatted" degrees of freedom – either on system "boundaries" or throughout – to have a specified

time history. Such a thermal constraint can be introduced in a variety of ways, but unique among them is the "Least Constraint", that which minimizes Gauss' constraint function, $\sum F_c^2/(2m)$. Consider applying an isokinetic constraint to the simplest possible system, with a separable energy:

$$E = \mathcal{H}(\{q, p\}) = K(\{p\}) + \Phi(\{q\}).$$

The constraint forces which fix the total kinetic energy of such a system, $K \equiv K_o$, and which simultaneously satisfy Gauss' Principle of Least Constraint, are linear in the momenta:

$$\{\dot{p} = F + F_c \overset{\text{Gauss}}{\rightarrow} F_c = -\zeta p\}.$$

By insisting that the time derivative of the isokinetic constraint be zero, we can identify the friction coefficient ζ:

$$0 \equiv \dot{K} = \sum (p/m) \cdot \dot{p} = \sum (p/m) \cdot (F - \zeta p) \;\Rightarrow\; \zeta = \sum (p/m) \cdot F \Big/ \sum (p^2/m) \equiv -\dot{\Phi}/2K_0.$$

The same constraint forces also follow from Hamilton's Principle of Least Action [17]. In terms of the accelerations $\{\ddot{r}\}$ the equations of motion for all the degrees of freedom thermostatted in this way have the form

$$\{\ddot{r}_{\mathcal{H}} = (F/m) - \zeta_{\mathcal{H}} \dot{r}\}, \quad \zeta_{\mathcal{H}} \equiv \sum \dot{r} \cdot F \Big/ \sum m\dot{r}^2 = -\dot{\Phi}/2K_0,$$

where the sums include all thermostatted degrees of freedom. A third route to these same constraint forces, the one which was first used in computations, is to rescale all thermostatted velocities by a common scaling factor,

$$v \rightarrow v \sqrt{2K_0 \Big/ \sum m v^2}.$$

Nosé developed a form of mechanics specifically designed to reproduce Gibbs' canonical distribution,

$$f_{\text{Gibbs}}(q, p) \propto e^{-\mathcal{H}/kT} = e^{-(K+\Phi)/kT}.$$

His derivation of the motion equations was unnecessarily complicated. There are two simpler routes, one discovered by Holian and the other by Dettmann [31]. Holian asked the question, "What friction coefficient ζ, with $\{\dot{p} = F - \zeta p\}$, is consistent with Gibbs' distribution?" The answer is

$$\zeta_{\text{NH}} \equiv \int_0^t [(K/K_0) - 1]\, dt'/\tau^2, \quad K_0 \equiv DNkT/2.$$

The equations of motion which result are identical to the "Nosé–Hoover" form of Nosé's equations. To establish the consistency of these equations with Gibbs' canonical distribution, it is only necessary to substitute the motion equations:

$$\{\dot{q} = p/m;\, \dot{p} = -\nabla\Phi - \zeta p\}, \quad \dot{\zeta} = [(K/K_0) - 1]/\tau^2,$$

into the phase-space continuity equation, using a generalized canonical phase-space probability density,

$$f(q, p, \zeta) \propto f_{\text{Gibbs}}(q, p) e^{-ND\tau^2\zeta^2/2}.$$

The result,

$$\begin{aligned} \partial f/\partial t &\equiv -\sum \partial(\dot{q} f)/\partial q - \sum \partial(\dot{p} f)/\partial p - \partial(\dot{\zeta} f)/\partial \zeta \\ &\equiv -\sum \dot{q} \cdot \partial f/\partial q - \sum \dot{p} \cdot \partial f/\partial p - \sum \zeta f - \dot{\zeta}\, \partial f/\partial \zeta \equiv 0, \end{aligned}$$

establishes the validity of Nosé's approach, and Holian's, directly. In the limit that τ approaches zero, the Nosé–Hoover trajectories approach those from Hamilton's principle.

Dettmann pointed out that exactly the same equations for the accelerations $\{\ddot{q}\}$ follow from an ad hoc Hamiltonian:

$$\mathcal{H}_D = \sum (p^2/2ms) + s\Phi + sNDkT \ln s + (sP_s^2/2NDkT\tau^2) \equiv 0,$$

by straightforward differentiation:

$$\{\dot{q} \equiv +\partial\mathcal{H}/\partial p = (p/ms),\ \dot{p} \equiv -\partial\mathcal{H}/\partial q = sF\} \ \Rightarrow\ \{\ddot{q} = (F/m) - (P_s/NDkT\tau^2)\dot{q}\}.$$

Here the "momentum" P_s is, apart from a multiplicative constant, equal to the Nosé–Hoover friction coefficient ζ. Here ND is the number of thermostatted degrees of freedom and τ is an arbitrary relaxation time.

Note that in Dettmann's approach the momentum p is no longer equal to $m\dot{q}$. In a conventional $\{q, p\}$ phase space, with p taken to be equal to $m\dot{q}$, Nosé–Hoover trajectories require an extension of Liouville's Theorem. The theorem states that phase volume is conserved by Hamilton's equations of motion. With time-reversible heat transfer included, phase volume grows whenever the system absorbs heat, and diminishes whenever heat is extracted from it. In such cases, the phase-space continuity equation expresses the change of comoving phase volume $\otimes$ with time in terms of the friction coefficient ζ:

$$\mathrm{d} \ln f/\mathrm{d}t = -\mathrm{d} \ln \otimes/\mathrm{d}t \equiv \sum \zeta.$$

As before, the sum includes each thermostatted degree of freedom.

Phase volume changes can also be described in terms of the Lyapunov spectrum, $\{\lambda\}$, as is described in the contribution of Harald Posch to these Proceedings. In classical Hamiltonian mechanics the Lyapunov exponents exist in pairs $\{+\lambda, -\lambda\}$, with the symmetry following directly from the time-reversibility of the Hamiltonian equations of motion. Their sum vanishes, corresponding to the conservation of phase volume. It is paradoxical that this symmetry relation is lost for time-reversible thermostatted systems, except at equilibrium. In the purely Hamiltonian version of the approach pioneered by Gaspard and Nicolis [19], the symmetry of the Lyapunov exponents is broken in a different way, by introducing sinks, through which particles can leave the system, and leading to a long-time fractal distribution for the remaining particles.

In either case, away from equilibrium, the sum of the Lyapunov exponents is typically negative, not zero. When thermostats are used, there is an intimate connection between the time-averaged friction coefficients and the Lyapunov exponents:

$$\sum \langle \zeta \rangle \equiv -\sum \lambda.$$

In the case of a nonequilibrium steady state this result suggests that $\ln f$ diverges to ∞ while the comoving phase volume $\otimes$ approaches zero. Numerical simulations with very simple systems showed that this interpretation is correct. The phase-space distribution typically "collapses" ($f \to \infty$; $\otimes \to 0$) onto an ergodic multifractal attractor, like those from the time-reversible map shown in Fig. 3. Fig. 4 shows a typical multifractal phase-space distribution found for the first such system to be investigated, the "Galton Board" or "Lorentz Gas", a thermostatted mass driven through a regular periodic lattice of hard elastic disks by a constant external field [12].

Several research groups have used this popular example of time-reversible thermostatted fractal dissipation to launch more-ambitious treatments of nonequilibrium systems [27,32–34]. The underlying goal of all this work is "understanding", uncovering and formalizing useful general principles for dealing with nonequilibrium systems. So far the new approaches have illustrated formal results and conjectures with very simple models, such as maps and the Lorentz Gas. We can expect further numerical verification of these ideas in the near future, but, due to the numerical difficulties involved, probably only for very small, idealized systems.

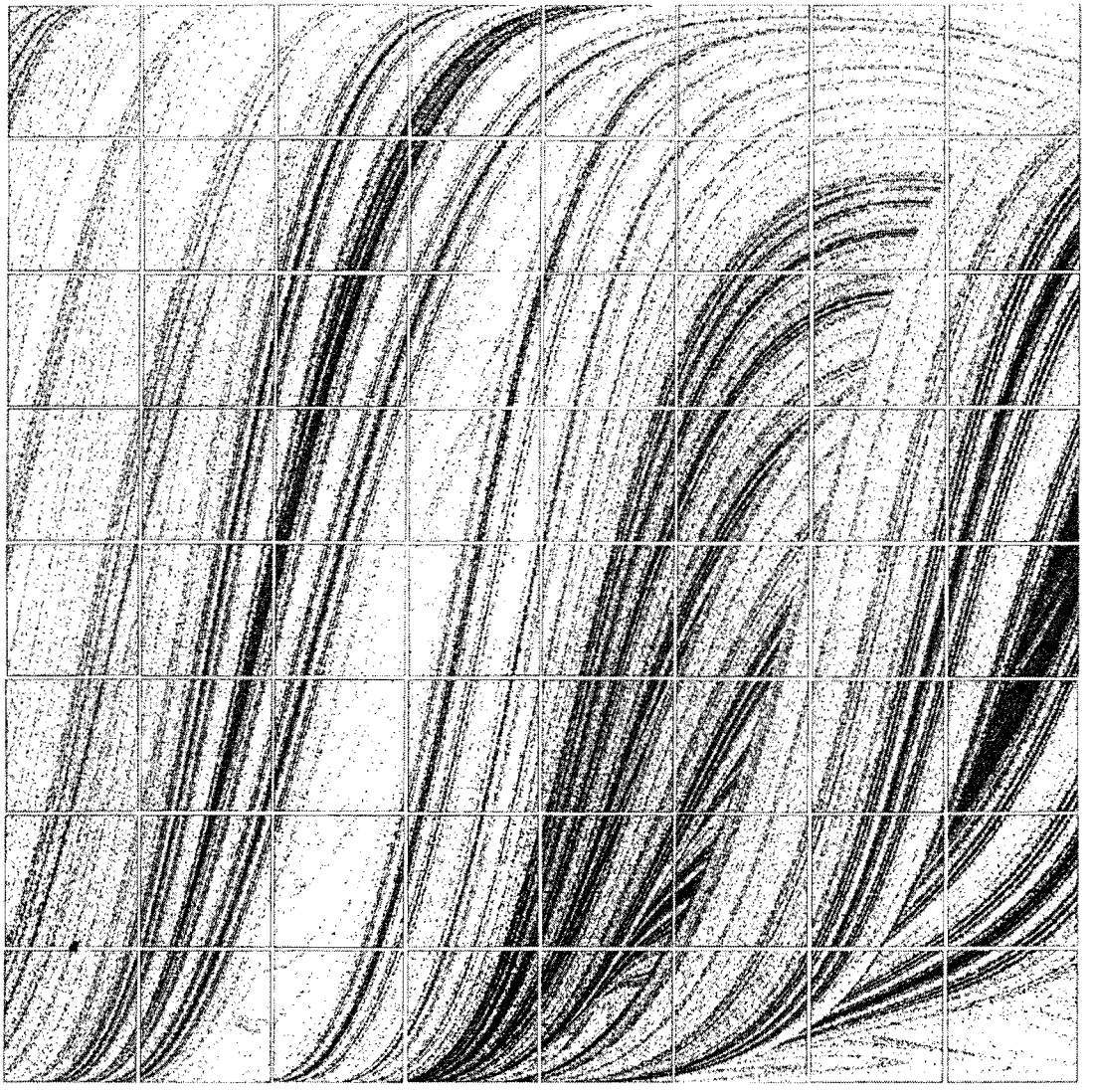

Fig. 4. A typical multifractal strange attractor describing the dynamics of a nonlinear one-body problem, the Galton Board. See [12] and the contribution of Posch and Dellago to these Proceedings.

The Galton–Board – Lorentz-Gas system is the model described and treated, by Posch and Dellago, in their contribution to this workshop. Numerous more-complex systems have been investigated, with up to several hundred degrees of freedom, and these studies have fully confirmed the conclusions based on thermomechanical simulations of simple few-body models [35]. In every case the phase-space probability density collapses onto a zero-volume strange attractor, characterized by a Lyapunov spectrum with a negative sum. The rate at which the phase volume decreases is given exactly by the entropy change of the external reservoirs, as represented by the reversible friction coefficients $\{\zeta\}$. If the process is periodic in the time, or stationary, so that global Lyapunov exponents $\{\lambda\}$ can be defined, then these too are related to the time-averaged value of the reservoirs' entropy change $\dot{S}$:

$$\sum \langle \dot{S}_{\text{external}}/k \rangle \equiv \langle -\mathrm{d} \ln \otimes / \mathrm{d}t \rangle \equiv -\sum \lambda \equiv \sum \langle \zeta \rangle > 0.$$

It seems paradoxical that this unidirectional collapse, onto a fractal attractor, is achieved with *time-reversible* motion equations. Take, for instance, Hamilton's form for the constrained least-action equations of motion:

$$\{\ddot{r}_{\rm H} = (F/m) - \zeta_{\rm H}\dot{r}\}.$$

Both $\{\ddot{r}\}$ and $\{F\}$ must be unchanged in any time-reversed trajectory satisfying these equations. On the other hand, both the friction coefficients $\{\zeta_{\rm H}\}$ and the velocities $\{\dot{r}\}$ do change sign, so that exactly the same motion equations hold, both forward and backward in time. They are *time reversible*. Nevertheless, just as in the examples using simple reversible maps, numerical solutions of these equations always satisfy, in a time-averaged sense, the Second Law of Thermodynamics, with mechanical work inexorably converted to heat.

The purely Hamiltonian simulations, with particle sinks, though not at all paradoxical, likewise lead to fractal distributions, despite the absence of thermostats. In either time direction particles with strongly correlated coordinates and momenta are lost from the initial phase-space distribution. This loss leads to an averaged long-time entropy production, for open systems, which can be written as a sum of two parts [21]:

$$\dot{S}/k = -\sum \lambda + \kappa, \quad \kappa = -\langle {\rm d}\ln N/{\rm d}t\rangle.$$

The summed Lyapunov exponents can result from coarse graining, which contributes an entropy gain equal to that which would be required for equivalent thermostatted heat losses. The remaining entropy production rate, κ, is special to open systems, and describes the shrinking of phase density due to particle losses. A recent lucid review of these ideas [21] should be consulted for details, including an example calculation and additional references.

The details of the unidirectional behavior found with thermostats can depend upon the numerical solution technique. A movie of the numerical solution, projected backward, shows a violation of the Second Law of Thermodynamics and makes no sense. It is noteworthy that the movie can be continued backward only to the initial state of the motion – the first frame. Any attempt to integrate the equations of motion further backward in time, beyond the initial condition, by replacing Δt with $-\Delta t$ produces a solution indistinguishable, on the average, from the solution forward in time.

Why is the trajectory obtained by changing the signs of all the velocities both silly and wrong? The answer lies in the Lyapunov instability of any such reversed trajectory. This instability necessarily arises because the phase-space strange attractor is far more stable than is its time-reversed image, the strange repellor, upon which f is a decreasing function of time so that the comoving phase volume $\otimes$ would have to increase. This asymmetry provides a basis for the analysis of phase-space distributions through unstable periodic orbits [29].

Evidently there can be no analog of the time-symmetric integer leapfrog algorithm for thermostatted systems. Whenever the phase volume undergoes change, in a space with integer coordinates, a faithful mapping must be locally either one-to-many or many-to-one. Neither mapping can be reversed. Except in very special circumstances – good random-number generators, for instance – any time-periodic mapping, reversible or not, cannot be simultaneously ergodic.

4. Least constraint and least action

Because Hamilton's Principle of Least Action provides the most general basis for the simulation of both equilibrium and nonequilibrium systems, it is worthwhile to apply the Principle to thermomechanics. If we express a many-body phase-space trajectory as a time series of coordinate sets $\{q(t)\}$ at discrete times $\{n\Delta t\}$, and joined by linear interpolation, we can approach the problem of minimizing the action integral,

$$\delta \int \mathcal{L}\,{\rm d}t \equiv \delta \int K\,{\rm d}t - \delta \int \Phi\,{\rm d}t = 0,$$

by a sequence of variations of the individual coordinates $\{q\}$. For a typical trajectory segment linking fixed coordinates at the times $-\Delta t$ and $+\Delta t$ through variable coordinates at the intermediate time, 0, the variation of the kinetic-energy integral – with the mass held fixed – is

$$\delta \int K \, \mathrm{d}t = \Delta t \sum m \delta q_0 \cdot [(q_0 - q_-) + (q_0 - q_+)]/\Delta t^2$$

and the variation of the potential-energy integral is

$$-\delta \int \Phi \, \mathrm{d}t = \Delta t \sum \delta q_0 \cdot F_0.$$

In the absence of any thermal constraint, setting the total variation equal to zero provides the Størmer–Verlet "leapfrog" algorithm [36].

If it is additionally required that the kinetic energy be kept constant, with

$$K_0 = \sum m(q_0 - q_-)^2/2\Delta t^2 = \sum m(q_+ - q_0)^2/2\Delta t^2,$$

this requirement can be satisfied with a single constraint, imposed by a Lagrange multiplier λ_1:

$$\lambda_1 \rightarrow \sum m \delta q_0 \cdot (q_0 - q_-) \equiv \sum m \delta q_0 \cdot (q_0 - q_+).$$

The equivalent differential equations of motion are obtained by letting the timestep Δt vanish. They are identical to those which follow from Gauss' Principle of Least Constraint:

$$\{\ddot{q} = (F/m) - \lambda_1 \dot{q}\}, \quad \lambda_1 = \zeta = \sum \dot{q} \cdot (F/m) \Big/ \sum \dot{q}^2.$$

It is worthwhile to consider briefly a more-elaborate approach which uses two separate multipliers to impose two separate isokinetic constraints:

$$\lambda_1 \rightarrow \sum m \delta q_0 \cdot (q_0 - q_-) \equiv 0, \qquad \lambda_2 \rightarrow \sum m \delta q_0 \cdot (q_0 - q_+) \equiv 0.$$

In this case, it appears that the limiting differential equations of motion would include two separate Lagrange multipliers, corresponding to the sum and difference of λ_1 and λ_2:

$$\{\ddot{q} = (F/m) - \zeta_1 \dot{q} - \zeta_2 \ddot{q}\}.$$

Just as before, the kinetic energy is constant, provided that the first multiplier, ζ_1, is identified with the Gaussian friction coefficient. The remaining multiplier, ζ_2, is indeterminate. There are three different grounds for discarding it. First, one could argue that mechanics should provide accelerations which are functions of coordinates and velocities alone. One could also discard the additional multiplier on the ground of simplicity – this approach is equivalent to applying Gauss' Principle of Least Constraint. Finally, any isokinetic multiplier should be a homogeneous function of the timestep Δt. It might well be interesting to shed further light on the subtle numerical mathematics of constrained isokinetic trajectories, by studying these for the simplest tractable case, the Galton Board [12].

5. Example problems

The mechanics problems which interest me are "chaotic". They have solutions which are "Lyapunov unstable". This chaotic instability means that small perturbations in the initial coordinates and momenta, $\{\delta q, \delta p\}$, will grow exponentially in the time, as $e^{\lambda t}$. This "sensitive dependence on initial conditions" typifies chaotic systems and, as we

shall see, affects the physical and computational, as opposed to mathematical, time reversibility of nonequilibrium systems. Many examples of chaotic problems are now known. The simplest describe nonequilibrium steady states of ordinary Newtonian or Hamiltonian dynamical systems, driven from equilibrium by explicit interactions with heat reservoirs, or, by the presence of particle sources and sinks.

The first problem to be investigated, the "Galton Board", which produces fractals of the type shown in Fig. 4, has all the generic characteristics of time-reversible nonequilibrium simulations – ergodicity, negative Lyapunov-exponent sum, a multifractal repellor–attractor pair in phase space. This problem has been more intensively investigated than any other, particularly in the last four years [29,32–34,37]. See again the contribution of Posch and Dellago to the Proceedings of this workshop.

More complicated many-body flows, driven by velocity or temperature differences, have shown that the Lyapunov spectra are relatively insensitive to the details of the nonequilibrium flows. Nevertheless, an investigation of the Rayleigh–Bénard flow shows that both the macroscopic and the microscopic instabilities can be seen in the Lyapunov spectrum [38]. Though full Lyapunov spectra cannot yet be obtained for systems with more than several hundred degrees of freedom, the same nonequilibrium simulation techniques have been applied to systems with millions of degrees of freedom [39]. The current size record, for a simulation of useful length, is about 10^8 [7]. See again Fig. 2.

6. Summary

Though time reversibility would seem to preclude strictly irreversible behavior, the introduction of thermal and particle sources and sinks into mechanics provides multifractal explanations of this paradox. Some explanations seem to be more complex than the phenomena they seek to explain, but, overall, there is an exciting sense of challenge and progress, both in the theoretical and the applied efforts. Though many of the details and interconnections are still missing, as is also an understanding of the importance of fluctuations, the near future appears much more promising than was the case just 15 years ago. The techniques suggested by the simple thermomechanical extensions of classical mechanics have also led to parallel discoveries in particle methods for solving continuum problems [23].

7. Remaining puzzles

I would be grateful for a better understanding of the fundamental paradox that time-reversible differential equations can lead to irreversible dissipative behavior. This paradox was noted by Illner and Neunzert [2]. It can be illustrated by the simple maps of Fig. 3, by the even simpler rotated generalized Baker's Map, by the more-elaborate chains of coupled maps studied by Gaspard [27], as well as through the analysis of simple model systems, both open [21] and closed [12]. So, good examples abound.

Perhaps the paradox linking reversibility and irreversibility is more subtle than the usual distinction between rational and irrational numbers, and includes the same fundamental mathematical difficulty, that which underlies the Banach–Tarski paradox. French illustrated that paradox [40] by constructing a one-to-one mapping, carrying a single point set into *two* disjoint sets, congruent with the original. French constructed two oranges from one in this way. Such "paradoxical decompositions" and constructions have the same peculiar nature as do "steady" phase-space flows with continuously shrinking distributions. It would be useful for a kind-hearted and knowledgable mathematician or physicist to express his understanding of such paradoxical behavior in more accessible terms.

Does finite-precision arithmetic have any important effect on the properties obtained from simulations? The difference between having many very long periodic orbits, with fixed precision, and the idealized inaccessible

ergodic chaotic orbit implied by classical continuum-based mechanics is reminiscent of the distinctions between quantum chaos (if any) and classical chaos. In nature h appears to be relatively larger than the corresponding precision limits of our simulations. Are there any interesting situations in classical simulation for which reasonable finite precision (say 12 digits) is inadequate?

So far the only nonequilibrium thermostats which have been linked to classical mechanics (through Hamilton's Principle of Least Action) are the isoenergetic or isokinetic ones, which follow also from Gauss' Principle of Least Constraint. The Nosé–Hoover thermostat, which is more elegant and flexible at equilibrium, has so far resisted every effort to derive it for nonequilibrium systems. Can Dettmann's single-temperature equilibrium derivation, or perhaps some entirely different approach, be used to link nonequilibrium Nosé–Hoover mechanics to conventional mechanics?

In one dimension, time-reversible maps lead nowhere. Time-reversible chaotic maps require at least a two-dimensional space. Chaotic flows require three dimensions. Two-dimensional maps can reproduce most properties associated with nonequilibrium systems, as we have seen. There appear to be problems simpler than maps, but perhaps conceptually just as fruitful, involving the manipulation of digit strings. Is there an analog of simple shear for digit strings? To what extent does the paradoxical coexistence of reversibility and irreversibility persist in simple models of this kind? On the other hand, the "simplest" continuous system which avoids periodic boundaries and discontinuities requires a coordinate, a reservoir variable, and a momentum, with the time development taking place in a three-dimensional phase space. Such a simple system can give rise to extremely complex strange attractors [41]. Increasing the dimensionality to 4, by adding another reservoir variable, leads to a much simpler, ergodic attractor. Are there simple topological reasons for this marked dependence of complexity and structure on dimensionality? To what extent can we claim to "understand" nonequilibrium flows on the basis of two-dimensional maps?

How can we characterize the multifractal distributions which characterize these nonequilibrium flows? The Lyapunov spectra seem to have less structure than do the frequency spectra of solid state physics. The spectrum $f(\alpha)$ seems to me to be a particularly uninteresting "featureless curve". Approximate representations of fractals as sums of periodic orbits are cumbersome in the extreme. Are there some computational ways, both "interesting" and "practical", to characterize all these attractors? These representations should reflect the dramatic differences in visual appearance of the few simple two- and three-dimensional cases accessible with computer graphics.

The Second Law of Thermodynamics is perfectly general, and applies to all nonequilibrium systems. On the other hand, it seems to be necessary to examine each individual special case to make predictions of constitutive behavior. Are there any guidelines or general rules, analogous to Gibbs' ensembles, which would simplify the computational task of predicting nonequilibrium behavior?

If any one of these questions can be answered, or rephrased so as to make an existing answer "interesting", and accessible for mathematicians and physicists alike, I shall consider this workshop to have been a huge success.

Acknowledgements

I began the work reported here with the help of Bill Ashurst and Bill Moran. It has continued to develop over the last 20 years. Denis Evans was an early source of inspiration for me. During the past decade I have particularly enjoyed collaborating with Harald Posch and Brad Holian. These men have contributed greatly to my views on reversibility in mechanics and provided useful advice, as has also Jeroen Lamb, for this review. In addition, a recent student, Oyeon Kum, carried out two separate projects which helped elucidate the reversibility of the smooth particle continuum equations and the ergodic time-reversible maps mentioned in the text. It was Bill Vance who pointed out to me that the rotated Baker Map satisfies the physicist's usual picture of time reversibility.

References

[1] J.A.G. Roberts and G.R.W. Quispel, Chaos and time-reversal symmetry, order and chaos in reversible dynamical systems, Phys. Rep. 63 (1992) 216.
[2] R. Illner and H. Neunzert, The concept of irreversibility in the kinetic theory of gases, Transport Theory Stat. Phys. 16 (1987) 89.
[3] R. Feynman, The Character of Physical Law (Modern Library, New York, 1994).
[4] B.J. Alder and T.E. Wainwright, Molecules in motion, Sci. Amer. 201 (4) (1959) 113.
[5] D. Levesque, L. Verlet and J. Kürkijarvi, Computer experiments on classical fluids. IV. Transport properties and time-correlation functions of the Lennard–Jones liquid near its triple point, Phys. Rev. A 7 (1973) 1690.
[6] B.L. Holian, W.G. Hoover, B. Moran and G.K. Straub, Shockwave structure via nonequilibrium molecular dynamics and Navier–Stokes continuum mechanics, Phys. Rev. A 22 (1980) 2798.
[7] B.L. Holian, P.S. Lomdahl and S.Y. Zhou, Fracture simulations via large-scale noncquilibrium molecular dynamics, Physica A 240 (1997) 340.
[8] J. Koplik and J.R. Banavar, Continuum deductions from molecular dynamics, Ann. Rev. Fluid Mechanics 27 (1995) 257.
[9] S. Nosé, Constant temperature molecular dynamics methods, Prog. Theor. Phys. (Supplement) 103 (1991) 1.
[10] W.G. Hoover and B.L. Holian, Kinetic moments method for the canonical ensemble, Phys. Lett. A 211 (1996) 253.
[11] D.J. Evans and B.L. Holian, The Nosé–Hoover thermostat, J. Chem. Phys. 83 (1985) 4069.
[12] B. Moran, W.G. Hoover and S. Bestiale, Diffusion in a periodic Lorentz gas, J. Stat. Phys. 48 (1987) 709.
[13] W.G. Hoover, H.A. Posch, B.L. Holian, M.J. Gillan, M. Mareschal and C. Massobrio, Dissipative Irreversibility from Nosé's Reversible Mechanics, Molecular Simulation 1 (1987) 79.
[14] W.G. Hoover, Computational Statistical Mechanics (Elsevier, Amsterdam, 1991).
[15] D.J. Evans and G.P. Morriss, Statistical Mechanics of Nonequilibrium Liquids (Academic Press, New York, 1990).
[16] B.L. Holian, W.G. Hoover and H.A. Posch, Resolution of Loschmidt's paradox: The origin of irreversible behavior in reversible atomistic dynamics, Phys. Rev. Lett. 59 (1987) 10.
[17] W.G. Hoover, Temperature, least action, and Lagrangian mechanics, Phys. Lett. A 204 (1995) 133.
[18] C.P. Dettmann and G.P. Morriss, Hamiltonian formulation of the Gaussian isokinetic thermostat, Phys. Rev. E 54 (1996) 2495; and references therein.
[19] P. Gaspard and G. Nicolis, Transport properties, Lyapunov exponents, and entropy per unit time, Phys. Rev. Lett. 65 (1990) 1693.
[20] P. Gaspard and J.R. Dorfman, Chaotic scattering theory, thermodynamic formalism, and transport coefficients, Phys. Rev. E 52 (1995) 3525.
[21] W. Breymann, T. Tél and J. Vollmer, Entropy production for open dynamical systems, Phys. Rev. Lett. 77 (1996) 2945; T. Tél, J. Vollmer and W. Breymann, Transient chaos: The origin of transport in driven systems, Europhys. Lett. 35 (1996) 659.
[22] B.L. Holian, G. Ciccotti, W.G. Hoover, B. Moran and H.A. Posch, Nonlinear response theory for Time-Independent fields: Consequences of the fractal nonequilibrium distribution function, Phys. Rev. A 39 (1989) 5414.
[23] D. Levesque and L. Verlet, Molecular dynamics and time reversibility, J. Stat. Phys. 72 (1993) 519; O. Kum and W.G. Hoover, Time-reversible continuum mechanics, J. Stat. Phys. 76 (1994) 1075.
[24] S.G. Brush, Kinetic Theory: 1, The Nature of Gases and of Heat; 2, Irreversible Processes; and 3, The Chapman–Enskog Solution of the Transport Equation for Moderately Dense Gases (Pergamon, Oxford, 1965, 1966, and 1972).
[25] N.G. van Kampen, The case against linear response theory, Phys. Norvegica 5 (1971) 279.
[26] W.G. Hoover, O. Kum and H.A. Posch, Time-reversible dissipative ergodic maps, Phys. Rev. E 53 (1996) 2123.
[27] P. Gaspard, Diffusion, effusion, and chaotic scattering: An exactly solvable Liouvillian dynamics, J. Stat. Phys. 68 (1992) 673; P. Gaspard, "What is the role of chaotic scattering in irreversible processes?, Chaos 3 (1993) 427.
[28] W.G. Hoover, B.L. Holian and H.A. Posch, Comment I on possible experiment to check the reality of a nonequilibrium temperature, Phys. Rev. E 48 (1993) 3196.
[29] W.N. Vance, Unstable periodic orbits and transport properties of nonequilibrium steady states, Phys. Rev. Lett. 67 (1992) 1356; G.P. Morriss, L. Rondoni and E.G.D. Cohen, A dynamical partition function for the Lorentz gas, J. Stat. Phys. 80 (1995) 35; L. Rondoni and G.P. Morriss, Stationary nonequilibrium ensembles for thermostatted systems, Phys. Rev. E 53 (1996) 2143.
[30] P.W. Bridgman, A physicist's second reaction to mengenlehre, Scripta Math. 2 (1934) 101.
[31] C.P. Dettmann, Private conversations at CECAM Euroconf. (Lyon, France, 1996); C.P. Dettmann and G.P. Morriss, Hamiltonian reformulation and pairing of Lyapunov exponents for Nosé–Hoover dynamics, Phys. Rev. E 55 (1997) 3693.
[32] N.I. Chernov, G.L. Eyink, J.L. Lebowitz and Y.G. Sinai, Derivation of Ohm's law in a deterministic mechanical model, Phys. Rev. Lett. 70 (1993) 2209.
[33] E.G.D. Cohen, Transport coefficients and Lyapunov exponents, Physica A 213 (1995) 293.
[34] G. Gallavotti, Chaotic hypothesis: Onsager reciprocity and fluctuation–Dissipation theorem, J. Stat. Phys. 84 (1996) 899.
[35] H.A. Posch and W.G. Hoover, Equilibrium and nonequilibrium Lyapunov spectra for dense fluids and solids, Phys. Rev. A 39 (1989) 2175.
[36] R.E. Gillilan and K.R. Wilson, Shadowing, rare events, and rubber bands. A variational Verlet algorithm for molecular dynamics, J. Chem. Phys. 97 (1992) 1757.

[37] H. van Beijeren, J.R. Dorfman, E.G.D. Cohen, H.A. Posch and Ch. Dellago, Lyapunov exponents from kinetic theory for a dilute, field-driven Lorentz gas, Phys. Rev. Lett. 77 (1996) 1974.
[38] T. Watanabe and H. Kaburaki, Increase in chaotic motions of atoms in a large-scale self-organized motion, Phys. Rev. E 54 (1996) 1504.
[39] J.S. Kallman, W.G. Hoover, C.G. Hoover, A.J. De Groot, S.M. Lee and F. Wooten, Molecular dynamics of silicon indentation, Phys. Rev. B 47 (1993) 7705.
[40] R.M. French, The Banach–Tarski Theorem, Math. Intell. 10 (4) (1988) 21.
[41] H.A. Posch and Wm.G. Hoover, Time-reversible dissipative attractors in three and four phase-space dimensions, Phys. Rev. E 55 (1997) 6803.