Journal of Statistical Physics, Vol. 109, Nos. 3/4, November 2002 (© 2002)

Comments on the Entropy of Nonequilibrium Steady States

Denis J. Evans[1] **and Lamberto Rondoni**[2]

Received September 28, 2001; accepted June 24, 2002

We discuss the entropy of nonequilibrium steady states. We analyze the so-called spontaneous production of entropy in certain reversible deterministic nonequilibrium system, and its link with the collapse of such systems towards an attractor that is of lower dimension than the dimension of phase space. This means that in the steady state limit, the Gibbs entropy diverges to negative infinity. We argue that if the Gibbs entropy is expanded in a series involving 1, 2,... body terms, the divergence of the Gibbs entropy is manifest only in terms involving integrals whose dimension is higher than, approximately, the Kaplan–Yorke dimension of the steady state attractor. All the low order terms are finite and sum in the weak field limit to the local equilibrium entropy of linear irreversible thermodynamics.

KEY WORDS: Chaos; fractal; entropy; nonequilibrium steady state.

1. INTRODUCTION

Over the last twenty years a set of dynamical systems has been defined that realistically model classical many particle nonequilibrium steady states. These dynamical systems can be used to accurately calculate the transport properties of fluids, their design being based on a precise connection between response theory for *mechanical* perturbations and various exact statistical mechanical expressions, such as Green–Kubo relations, for *boundary driven thermal* transport processes.(1) These dynamical systems can also be used to study changes to the molecular structure and conformation

---

[1] Research School of Chemistry, Australian National University, Canberra 0200, Australia; and Physical and Chemical Properties Division, NIST, Boulder, Colorado 80305; e-mail: evans@rsc.anu.edu.au

[2] Mathematics Dept., Politecnico di Torino, C. Duca degli Abruzzi 24, 10129 Torino, Italy.

0022-4715/02/1100-0895/0 © 2002 Plenum Publishing Corporation

caused by the action of the dissipative nonequilibrium forces that prevent the systems from relaxing to equilibrium.(1) In these systems it is frequent for fictitious (i.e., unnatural) mechanical forces to replace subtle thermal forces that act on systems by virtue of boundary conditions (walls maintained at different temperatures or walls that are maintained in relative motion). Fictitious deterministic thermostats that may operate homogeneously throughout the system, or that may operate within boundary walls, remove the dissipative heat generated in the system by the application of the dissipative fields, thereby allowing the possibility of a nonequilibrium steady state. We refer to these systems as Non-Equilibrium Molecular Dynamic (NEMD) systems.

A considerable literature (cf. refs. 1, 2, and references therein) demonstrates that nonequilibrium statistical mechanics can be used to validate NEMD systems in the sense that these dynamical systems can be used to calculate (essentially exactly), the nonequilibrium properties of real thermal systems. This exactness is usually limited by the approximate nature of the potential functions used to model molecular interactions. In the linear response regime, close to equilibrium, given an assumed molecular potential function, the calculated properties can usually be shown to be exact. For example, thermostatted time correlation functions are equal at all times (to leading order in the particles number N) to the corresponding adiabatic equilibrium correlation functions (cf. Section 7 in ref. 2). All the standard Navier–Stokes transport processes for both single component(1) and multicomponent mixtures, and certain types of liquid crystals,(2) can be characterized essentially exactly in the linear response regime. Even for the nonlinear regime far from equilibrium, many algorithms have been developed that model real systems extremely accurately.(1)

Despite these successes, one feature of NEMD models has puzzled some theoreticians: the equations of motion for steady state NEMD systems cannot be usually derived from an Hamiltonian, and therefore phase space volumes are not preserved by the associated dynamics. At first, this looks in striking contrast with the behavior of real systems subject to the action of a real thermostat, which is fully Hamiltonian.

However, there are various facts that one should keep in mind. Suppose we have a Hamiltonian model of a nonequilibrium steady state system made of a small nonequilibrium system which is subject to an external field and is in contact with a large (many particles) thermostat. The combined subsystem and thermostat are Hamiltonian and therefore phase space conserving. In order to treat a nonequilibrium *steady state*, the number of degrees of freedom in the thermostat must be very much larger than those of the nonequilibrium subsystem. Only when this is the case can we be sure that the heating rate is small relative to the characteristic relaxation time scales of the subsystem and the subsystem can therefore be regarded as being effectively at constant temperature.

Now, if we project out the myriad of thermostatting degrees of freedom, and consider only the motion of the (relevant) degrees of freedom of the subsystem, then it is obvious that the remaining degrees of freedom will exhibit a motion which is not phase space conserving.[3] The major effect of the myriad of thermostatting degrees of freedom is to prevent the natural heating that would otherwise occur in the dissipative Hamiltonian subsystem. Thus, one expects that the motion of the projected subsystem will exhibit compression as the result of heat being absorbed by the thermostatting degrees of freedom.

It is of course very cumbersome to treat this large composite Hamiltonian system theoretically. The deterministic and time reversible thermostats employed in NEMD aim at accomplishing this "projecting out" of irrelevant thermostatting degrees of freedom in a concise and clearly defined manner. In NEMD, the equations of motion for the projected subspace are derived by applying Gauss' Principle of Least Constraint(1) to the relevant degrees of freedom, in order to mimic the effect on the subsystem of the thermostatting degrees of freedom. Thus, we apply Gauss' Principle to fix the temperature or the internal energy etc. of the relevant subsystem. The constraint force required to enforce the thermodynamic constraint has the smallest magnitude possible. In this sense the constraint exerts the least possible influence on the phase space trajectories in the relevant phase space. In fact, for holonomic constraints Gauss' Principle is equivalent to Hamilton's Principle of Least Action.

One expects that different representations of the same physical system should be possible, as is the case in equilibrium statistical mechanics, where one observes that, for instance, the canonical and microcanonical descriptions become equivalent in the thermodynamic limit. In fact, one expects the detailed structure of any (proper) thermostat should not to play a fundamental role in determining the properties of the system to which the thermostat is applied—especially in the limiting weak field regime close to equilibrium. This leads to the possibility of extending the theory of equivalence of ensembles, which is well developed in equilibrium statistical mechanics, to the case of near equilibrium systems. That this may actually be possible is indicated by various works published in the last decade or so (see, e.g., refs. 2–5). Then, one may argue that different models of a nonequilibrium system could be equally valid in the weak field limit. For

[3] Take a 2-dimensional phase space with some volume preserving dynamics which combines both stretching and compressing local directions. If we project out one of the phase space coordinates, then the dynamics does not preserve the volumes in the projected subspace.

instance, using stochastic boundary conditions and deterministic thermostats leads to different phase space distributions, but should not lead to different transport properties. Thermostatted dynamical systems are also validated by recent experimental results concerning entropy production rates.(6, 7)

In this context, it is striking that Gibbs entropy of the nonequilibrium steady state generated in the NEMD systems is not constant, like all other thermodynamic properties (energy, pressure, temperature, specific heat, etc.). The Gibbs entropy decreases in time towards negative infinity. It seems to us that this problem cannot be dismissed by simply arguing that the Gibbs entropy shouldn't be used at all in nonequilibrium statistical mechanics, or invoking the "non-physical" nature of NEMD systems. In fact, it is obvious that a diverging quantity such as the Gibbs entropy cannot be the physical entropy of any system, but the problem persists for situations arbitrarily close to equilibrium, and for all deterministic dissipative dynamics. Hence the question is: why does this quantity behave in such a peculiar way?

Unlike other thermodynamic properties the Gibbs entropy, S_G, of a system is not a simple phase average but is rather a functional of the N-particle phase space distribution function $f(\Gamma, t)$ which characterizes the time evolution of an initial ensemble $f(\Gamma, 0)$, where Γ is one point in the $2dN$ dimensional phase space Ω. Typically the initial ensemble of NEMD simulations is one of the classical equilibrium ensembles of statistical mechanics (microcanonical, canonical, grand canonical, etc.). Without at this stage being explicit about the initial ensemble, the Gibbs entropy of a nonequilibrium N-particle system at time t, in d_C Cartesian dimensions, is defined as:

$$S_G(t) = -k_B \int d\Gamma\, f(\Gamma, t) \ln[f(\Gamma, t)], \tag{1}$$

where k_B is Boltzmann's constant. The form of the equations of motion which define the dynamical system and which implicitly control the time evolution of all properties including the Gibbs entropy take the generic form,

$$\begin{aligned} \dot{\mathbf{q}}_i &= \frac{\mathbf{p}_i}{m} + \mathsf{C}_i(\Gamma) \cdot \mathbf{F}_e \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i(\mathbf{q}) + \mathsf{D}_i(\Gamma) \cdot \mathbf{F}_e - S_i \alpha(\Gamma)\, \mathbf{p}_i \end{aligned} \tag{2}$$

where $\mathbf{F}_e$ is the dissipative external field that couples to the system via the phase functions C, D, and α is a Lagrange multiplier used to constrain the value of a properly chosen dynamical quantity. The term α in Eq. (2) acts in a way to remove (or add) the heat generated (or absorbed) in the system by the action of the external field, hence the name "thermostat." It is convenient to employ a switch, S_i, which controls how many particles are employed as thermostats ($S_i = 1$ only for wall particles or $S_i = 1, \forall i$, defines the so-called homogenous thermostat$^{(8)}$). It is worth pointing out that as described, Eqs. (2) are time reversible and heat can be both absorbed and given out by the thermostat. However, in accord with the Second Law of Thermodynamics, in dissipative dynamics the ensemble averaged value of the thermostat multiplier is positive at all times$^{(4)}$ ($\langle \alpha(t) \rangle > 0, \forall t > 0$).

Obviously, one should not confuse a real thermostat composed of a very large (in principle, infinite) number of particles with the purely mathematical—albeit convenient—term α. In writing (2) it is assumed that the momenta $\mathbf{p}_i$ are peculiar (i.e., measured relative to the local streaming velocity of the fluid or wall). The thermostat multiplier may be chosen, for instance, to fix the internal energy of the system $H_0 \equiv \sum_{i:\, S_i = 0} [p_i^2/2m + 1/2 \sum_j \phi(q_{ij})]$, where we have assumed that particles i, j interact through a pair potential, ϕ, in which case we speak of ergostatted dynamics, or we can constrain the peculiar kinetic energy of the wall particles $K_W \equiv \sum_{S_i = 1} p_i^2/2m = dN_W k_B T_W/2$, with $N_W = \sum S_i$, in which case we speak of isothermal dynamics. The quantity T_W defined by this relation is called the *kinetic* temperature of the wall. For homogeneously thermostatted systems, T_W becomes the kinetic temperature of the whole system and N_W becomes just the number of particles N, in the whole system.

One can use the time reversible Liouville equation corresponding to (2) to calculate the change in the Gibbs entropy of an ensemble of systems representing a given nonequilibrium physical situation, and then check how this change is related to the expected behavior of the *real* (physical) entropy in that nonequilibrium situation. The expectation is that the *real* entropy of the nonequilibrium steady state is finite, and smaller than that of the equilibrium state at the same thermodynamic state point (of same N, V and energy). Indeed, if the entropy is viewed as a measure of the degree of disorder in the system associated with the microscopic motions, this disorder is lower in a nonequilibrium steady state than in the corresponding equilibrium state. However, if we differentiate (1) we see that, for an N-particle system,

$$\dot{S}_G(t) = -k_B \int d\Gamma \{1 + \ln[f(\Gamma, t)]\}\, \partial f(\Gamma, t)/\partial t$$

$$= -k_B \int d\Gamma \ln[f(\Gamma, t)]\, \partial f(\Gamma, t)/\partial t$$

$$= k_B \int d\Gamma \ln[f(\Gamma, t)] \, \partial/\partial\Gamma \cdot [\dot{\Gamma} f(\Gamma, t)]$$

$$= -k_B \int d\Gamma [\dot{\Gamma} \cdot \partial/\partial\Gamma] \, f(\Gamma, t)$$

$$= k_B \int d\Gamma \; f(\Gamma, t) \, \partial/\partial\Gamma \cdot \dot{\Gamma}$$

$$= -d_C N_W k_B \int d\Gamma \; f(\Gamma, t) \, \alpha(\Gamma) + O(1)$$

$$= -d_C N_W k_B \langle \alpha(t) \rangle \tag{3}$$

In deriving (3), it was assumed that all the boundary terms arising from the successive integration by parts, vanish. This is the case for *confined* systems, whose N-particle distribution function vanishes at and outside the container walls, and for distribution functions that decay sufficiently rapidly with increasing momenta. For systems under periodic boundary conditions, the distribution function can be defined by mapping the whole infinite phase space to a torus representing the fundamental domain (the periodic tile of the whole phase space), and this likewise forces spatial boundary terms to vanish, while the momentum boundary terms vanish for the same reason as for confined systems. In the last equality, the $O(1)$ term has been neglected, assuming that N_W is large.[4] We also note that if the deterministic thermostat is applied homogeneously, the only change in (3) is that the factor N_W changes to N.

Equation (3) shows that the rate of change of the Gibbs entropy of the entire system (including the walls) is related to the *ensemble* average $\langle \alpha(t) \rangle$, of the thermostat multiplier α, that is positive for all times t after the external dissipative field is applied.(9) In accord with the Laws of Thermodynamics, the ensemble average of the work performed on the system by the external field is positive and appears as *heat*, which is removed by the thermostat in order to maintain the system at a constant internal energy or kinetic temperature. Equation (3) then shows that, after the decay of initial transients, the rate of decrease of the entropy is constant, implying that the Gibbs entropy of a *steady state* system diverges to negative infinity! On the other hand, if there is no thermostat (e.g., in the absence of a driving field), the Liouville equation predicts that the Gibbs entropy of an arbitrary

---

[4] The assumption that N, N_W are large will be made throughout this paper. For example, we will refer to the ostensible dimension of phase space as $2d_C N$ rather than $2d_C N - 2d_C - 1$, etc.

Hamiltonian system subject to an external dissipative field, is constant! This result was known to Gibbs (1902).(10)

Consider the ergostatted dynamics, characterized by the equation

$$\dot{H}_0(\Gamma) \equiv -J(\Gamma)\, VF_e - 2K_W(\Gamma)\, \alpha(\Gamma) = 0, \tag{4}$$

where J is the dissipative flux due to F_e and V is the volume of the system. Equation (4) is a statement of the First Law of Thermodynamics for an ergostatted nonequilibrium system. The energy removed (or added) to the system by the ergostat must be balanced instantaneously by the work done on (or removed from) the system by the external dissipative field, F_e.

In analogy with conventional thermodynamics, it is natural to *define* the total rate of absorption (or emission) of entropy by the ergostat as,

$$\Sigma(t) \equiv 2K_W(\Gamma)\, \alpha(\Gamma)/T_W(t) = d_C N_W k_B \alpha(t) = -J(t)\, VF_e/T_W(t), \tag{5}$$

[i.e., it is the heat that flows to (or is removed from) the ergostat divided by its absolute temperature]. The entropy flowing into the ergostat results from a continuous generation of entropy in the dissipative system. The so-called entropy source strength $\sigma(\mathbf{r})$ is the rate of entropy production per unit time and per unit volume at a position $\mathbf{r}$, in the system. The entropy source strength can be shown, close to equilibrium, to be a product of the so-called (local) thermodynamic fluxes and forces, and in a steady state the entropy that is produced spontaneously throughout the system is convected and/or conducted to (or away from) the boundary thermostat in which it is absorbed (or released).

Comparing (3) and (5) we see that for ergostatted dynamics

$$\langle \Sigma(t) \rangle = -\dot{S}_G(t). \tag{6}$$

The ensemble averaged entropy absorbed (emitted) by the ergostat per unit time is equal and opposite to the rate of change of the Gibbs entropy of the entire system.

Our *definition* (5) invokes a kinetic measure of the time dependent thermodynamic temperature. There are many measures of the temperature that could be used.(11) Sufficiently close to equilibrium, where the condition of local thermodynamic equilibrium holds, each of these measures lead, as they do at equilibrium, to identical values for the estimated thermodynamic temperature. Further from equilibrium, (5) is a formal *definition* of the rate of absorption of entropy by the ergostat, since classical thermodynamics and linear irreversible thermodynamics say nothing on this subject beyond the strictly equilibrium and local equilibrium regimes.

For an ensemble of constant internal energy *nonequilibrium* systems, we have

$$S_G(t) + \int_0^t \langle \Sigma(s) \rangle \, ds = S_{eq}(0), \tag{7}$$

where it is always understood that the system is at equilibrium at $t = 0$, and $S_{eq}(0)$ is the corresponding initial equilibrium entropy.

Similarly to ergostatted systems, for isothermal dynamics we obtain:

$$\lim_{t \to \infty} \langle \bar{\Sigma}(t) \rangle = -\lim_{t \to \infty} \dot{\bar{S}}_G(t), \tag{8}$$

where the bar indicates time averaging for phase variables X, i.e., $\bar{X}(t) = \frac{1}{t} \int_0^t X(s) \, ds$.

2. ERGODICITY

Most of nonequilibrium statistical mechanics is based on the assumption that the systems being studied are *ergodic*. For example, the Chapman–Enskog solution of the Boltzmann equation is based on the tacit assumption of ergodicity. The response theory relations for Navier–Stokes transport processes such as shear flow show that if the shear rate $\gamma(t)$ takes the form of a Heaviside step function of time, $\gamma(t) = \Theta(t)\, \gamma$, then the time dependent response of an equilibrium ensemble of systems takes the form$^{(1)}$

$$\langle P_{xy}(t) \rangle = \frac{-V}{k_{\mathrm{B}} T} \int_0^t ds \langle P_{xy}(s)\, P_{xy}(0) \rangle \, \gamma \tag{9}$$

where P_{xy} is the off diagonal term of the pressure tensor, and the brackets $\langle \cdots \rangle$ denote the average over an ensemble that is at equilibrium at $t = 0$, but which has subsequently had an ergostatted shear applied $\gamma(t) = \Theta(t)\, \gamma$. In the zero shear rate limit, Eq. (9) reduces to the familiar Green–Kubo relation which also describes the ensemble averaged response, and the right hand side of (9) involves an *equilibrium* (i.e., $\gamma = 0$) time correlation function.

In the long time limit, the strain rate dependent shear viscosity

$$\eta(\gamma) \equiv \lim_{t \to \infty} \frac{\langle P_{xy}(t) \rangle}{\gamma} \tag{10}$$

is easily seen to be

$$\eta(\gamma) = \frac{-V}{k_{\mathrm{B}} T} \int_0^\infty ds \langle P_{xy}(s)\, P_{xy}(0) \rangle. \tag{11}$$

In practice, however, experimentalists usually substitute time averages for ensemble averages. They therefore view the systems they study as *ergodic*:

$$\lim_{t\to\infty} \frac{\langle P_{xy}(t)\rangle}{\gamma} = \lim_{t\to\infty} \frac{\bar{P}_{xy}(t)}{\gamma}. \tag{12}$$

The fact is that experimentally one usually does not measure transport coefficients as ensemble averages. Almost universally transport coefficients are measured as time averages. Arguably, the clearest indication of the ubiquity of nonequilibrium ergodicity, which justifies the use of relations such as (12), is that empirical data tabulations assume that transport coefficients are single valued functions of both the thermodynamic state point (N, V, T) and possibly the strength of the dissipative field. The tacit assumption of nonequilibrium ergodicity is so widespread that it is frequently forgotten that it is in fact an *assumption*.

The precise conditions for ergodicity are not known. However, systems that appear ergodic at equilibrium usually remain ergodic under mild perturbations, away from equilibrium.

Assuming ergodicity, Eq. (6) becomes,

$$\lim_{t\to\infty} \bar{\Sigma}(t) = -\lim_{t\to\infty} \dot{S}_G(t). \tag{13}$$

It is important to recall that the notion of ergodicity used here does not require that a single phase space trajectory approaches within an arbitrarily small tolerance every point in the ostensible phase space.[5] This notion only requires that under a given set of macroscopic conditions, time averages and ensemble averages should be equal.

3. LYAPUNOV EXPONENTS

Since the divergence of the vector field $\dot{\Gamma}$ is $d_C N_W \alpha + O(1)$, the formal solution of the Liouville equation

$$\frac{df(\Gamma(t))}{dt} = -f(\Gamma(t))\,\frac{\partial}{\partial\Gamma}\cdot\dot{\Gamma}((\Gamma(t)) \tag{14}$$

is given by

$$f(\Gamma(t), t) = f(\Gamma(0), 0)\exp\left[\int_0^t ds\, d_C N_W \alpha(s)\right]. \tag{15}$$

---

[5] In fact, quite the opposite may be true.

From the definition of Lyapunov exponents, it is then obvious that if the trajectory in phase space originates from an initial point $\Gamma(0)=\Gamma_0$ picked at random with respect to the equilibrium probability distribution, and since the distribution function is normalized, the long time average of the phase space compression factor is related to the sum of the local Lyapunov exponents, by

$$\lim_{t\to\infty} d_C N_W \bar{\alpha}(t;\Gamma_0) = -\sum_{i=1}^{2d_C N} \lambda_i(\Gamma_0). \tag{16}$$

Assuming ergodicity, local Lyapunov exponents $\{\lambda_i(\Gamma_0)\}$ take the same values except for a set of initial Γ_0 of zero probability; thus, substituting (16) into (3) we obtain:

$$\lim_{t\to\infty} \dot{S}_G(t) = k_{\mathrm{B}} \sum_{i=1}^{2d_C N} \lambda_i. \tag{17}$$

We can understand this relation in a more physically revealing way by noting that if the distribution function is approximately constant over the accessible phase space then,

$$S_G(t) \approx k_{\mathrm{B}} \ln[V_\Gamma(t)] = S_{\mathrm{B}}(t), \tag{18}$$

where $V_\Gamma(t)$ is the volume of the accessible phase space at time t, and S_{B} is the Boltzmann entropy. More precisely, we have $S_G(t)=k_{\mathrm{B}} \ln V_\Gamma(t)$ for the probability distributions that are constant over the accessible phase space of volume $V_\Gamma(t)$, (i.e., for those distributions defined by:

$$\mu_t(d\Gamma) = \begin{cases} d\Gamma/V_\Gamma(t) & \text{if} \quad \Gamma\in\Omega_t\subset\Omega \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where Ω_t is a (possibly time dependent) subset of the full or "ostensible," phase space Ω, and $d\Gamma$ is a volume element of Ω). Differentiating (18) with respect to time gives:

$$\dot{S}_G(t) \approx k_{\mathrm{B}} \frac{\dot{V}_\Gamma(t)}{V_\Gamma(t)} \tag{20}$$

which links the variation of the Gibbs entropy to the rate of decrease in the volume of accessible phase space. On the other hand, the time average rate of variation of $V_\Gamma(t)$ is, in ergodic systems, given by the sum of the Lyapunov exponents of the system, which leads to:

$$\lim_{t\to\infty} \dot{S}_G(t) \approx k_\mathrm{B} \sum_{i=1}^{2d_C N} \lambda_i. \tag{21}$$

This leads us to a consideration on the Kaplan–Yorke dimension D_KY,(12) defined by

$$D_\mathrm{KY} = N_\mathrm{KY} + \frac{\sum_{i=1}^{N_\mathrm{KY}} \lambda_i}{|\lambda_{N_\mathrm{KY}+1}|}, \tag{22}$$

where the Lyapunov exponents are arranged in decreasing order ($\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{2d_C N}$), and N_KY is the largest integer for which $\sum_{i=1}^{N_\mathrm{KY}} \lambda_i$ is positive. We can say that, D_KY estimates the "dimension" of that subset of the phase space whose volume does not decrease with time. In other words, since D_KY can be a noninteger, finite volume elements of integer dimension $m < D_\mathrm{KY}$, with sides aligned with the eigendirections associated with the first m Lyapunov exponents, $\lambda_1,\ldots, \lambda_m$, expand in time, while volume elements of integer dimension $m > D_\mathrm{KY}$, contract in time.

This implies that if we could compute what we call the *reduced* Gibbs entropy over the appropriate volume of dimension D_KY, we would find that the steady state value of such a quantity is finite and is a constant of motion. This is precisely the reason why S_G is a constant for Hamiltonian systems: D_KY is an integer that is equal to the dimension of the ostensible phase space ($\sum_{i=1}^{2d_C N} \lambda_i = 0$ for Hamiltonian systems), and the accessible phase space volume neither expands nor contracts. In the case where D_KY is not an integer but the ostensible phase space dimension is large,[6] the Lyapunov spectrum is expected to approach a continuum distribution,(13) leading to $N_\mathrm{KY} \approx D_\mathrm{KY}$. Hence in a nonequilibrium steady state, *certain* N_KY-dimensional volume elements are expected to remain approximately unaltered by the time evolution. The singularities of the full probability distribution function, which are due to the phase space contraction, do not manifest themselves on these N_KY-dimensional volume elements.

It is important to remember that for real thermodynamic systems the actual dimensional decrease is tiny. For 1 mole of liquid argon sheared at the maximum shear rate that is possible before the flow would become turbulent, the difference between the Kaplan–Yorke dimension and the ostensible phase space dimension is $O(1)$.(13) The relative dimensional

---

[6] Of order $O(N_\mathrm{A})$, $N_\mathrm{A} =$ Avogadro's Number $= O(10^{23})$.

decrease is JUST $(6N_A - D_{KY})/6N_A \approx 10^{-23}$! Consequently, it is relevant to consider the behavior of nonequilibrium systems in the weak field limit. Recently, an interesting relation has been found between transport coefficients and D_{KY}, in the weak field limit.(13) Combining this relation with our (17), we find an exact relation for the rate of change of the Gibbs entropy in terms of the Kaplan–Yorke dimension, in the weak field limit:

$$\lim_{t\to\infty} \dot{S}_G(t) = k_B \sum_{i=1}^{2d_C N} \lambda_i \approx -k_B \lambda_{\max}(F_e = 0)(2d_C N - D_{KY}(F_e)). \tag{23}$$

The average rate of decrease in the steady state Gibbs entropy is for weak fields, proportional to the product of the maximal equilibrium Lyapunov exponent and the dimensional decrease of the steady state attractor from the ostensible phase space dimension.

We now see why the Gibbs entropy of a system keeps decreasing. If we start from an ensemble of equilibrium systems, and evolve them in time towards a nonequilibrium steady state, the volume of the phase space accessed by the ensemble collapses in time, and vanishes in the steady state limit. While in equilibrium a phase space subset can be assigned a positive probability only if it has a non-vanishing (ostensible) phase space volume, in the case of nonequilibrium steady states there are sets with positive probabilities but vanishing volumes. Since the Gibbs entropy is computed as an integral over the ostensible phase space, the steady state Gibbs entropy diverges towards negative infinity, as the system converges to the steady state.

From a mathematical point of view, this is expressed by the fact that the Gibbs entropy is defined only for phase space distributions which are absolutely continuous with respect to the Lebesgue measure on the ostensible phase space, i.e., distributions of the form $d\mu(\Gamma) = f(\Gamma)\, d\Gamma$ as used above, where f is the (integrable) density function. This is a consequence of the fact that the Gibbs entropy is not the average of a generic (e.g., smooth) phase variable but is, in a sense, the average of the logarithm of the distribution itself. Now, the dissipativity of NEMD dynamics in the presence of any (hence also arbitrarily small) nonvanishing field implies that the corresponding steady state distribution has no density, i.e., that the distribution is singular, and that the Gibbs entropy does not exist.

4. A STRIKING FACT

The classical subject of linear irreversible thermodynamics(14) is based on the assumption that in nonequilibrium states the local *nonequilibrium* thermodynamic entropy per particle is, in the weak field limit, the same as the *equilibrium* thermodynamic entropy per particle for an equilibrium system at the corresponding thermodynamic state point (i.e., local temperature and density). Our considerations on the divergence of the Gibbs entropy imply that in contradistinction to the situation at equilibrium, and regardless of how close a system may be to equilibrium, the Gibbs entropy of dissipative systems—such as the NEMD models—*cannot* be the nonequilibrium thermodynamic entropy as is used in linear irreversible thermodynamics. The striking fact is that there seems to be no smooth transition of the probability phase space distribution from its regular equilibrium form to its singular nonequilibrium steady state form.

Nevertheless, the entropy of a system prepared in an equilibrium state coincides with its Gibbs entropy, and linear irreversible thermodynamics is a *successful* macroscopic theory built around the assumption of local thermodynamic equilibrium and therefore of a finite nonequilibrium entropy. This suggests that a dynamical definition of a nonequilibrium entropy should be related to S_G, at least in the weak field regime, close to thermodynamic equilibrium.

Also, the abrupt loss of meaning undergone by S_G in the transition from the equilibrium state to *any* nonequilibrium steady state appears to be at odds with the analysis of the previous section, which shows that only small dimensional changes affect the attractors of Gaussian systems subject to small field variations around zero. This is further evidenced in the portraits of the phase space distributions of such simple (and low dimensional) systems as the nonequilibrium Lorentz gas.(15)

For instance, consider the equations of motion:

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p} \\ \dot{\mathbf{p}} &= \mathbf{E} \cdot (\mathbf{1} - \mathbf{pp}/p^2) \end{aligned} \tag{24}$$

for one point particle of mass 1, moving with constant speed ($p^2 = p_x^2 + p_y^2 = 1$), which is elastically scattered by disks placed on a regular triangular lattice tiling the 2-dimensional plane, and subjected to an external field $\mathbf{E} = (E, 0)$. Let us take disks of radius 1, and let the closest interdisk distance be $w = 0.236$, as in ref. 16. We note that the ostensible phase space dimension for these equations of motion is 3. This is already too high a dimension to conveniently visualize, but there is a direct correspondence between the motion in this 3-D hypersurface and the dynamics of the Poincaré map obtained by stroboscopically following the motion of the particle from collision to collision. In this case one looks only at the time evolution of the pair of coordinates (ϕ, η) where $\phi \in [0, 2\pi] \equiv I_\phi$ is the angle with respect to the x-direction which individuates the collision point on the surface of a scatterer, $\eta = \sin(\theta - \phi) \in [-1, 1] \equiv I_\eta$ and $\theta \in [-\pi, \pi]$ is the angle with respect to the x-direction of the outgoing velocity right after a collision. The coordinate η is the component of $\mathbf{p}$ along the line tangent to a scatterer at the collision point. The dynamics in the (ϕ, η)-plane is then easy to visualize and carries all the nontrivial information of the dynamics.

The equilibrium ($\mathbf{E} = 0$) invariant probability distribution of the coordinates (ϕ, η) is uniform in $\Omega = I_\phi \times I_\eta$, while the presence of an external field leads to fractal distributions which are more and more evident as E grows. However, despite the fact that the linear regime for one such model of Lorentz gas can be estimated to be bound by $|E| < 10^{-6}$,(17) so that larger fields should reveal themselves quite clearly in the phase space portraits, our Figs. 1 and 2 show that this is not the case below rather large fields, i.e., fields of order $O(10^{-1})$. This is further confirmed by the study of the generalized dimensions of the attractors of the nonequilibrium Lorentz gas,(18–20) which shows that the coarser generalized dimensions remain equal to 2 in the presence of small fields, while the finer generalized dimensions vary smoothly, decreasing as the field grows. In other words, the invariant probability distribution of the nonequilibrium Lorentz gas does not appear to undergo any abrupt change from the equilibrium to the near equilibrium steady states, at least on a level of description that seems appropriate for physical observations. The same is expected to be the case for the many particle systems described by the more general equations of motion (2), albeit not directly testable with present day technology.

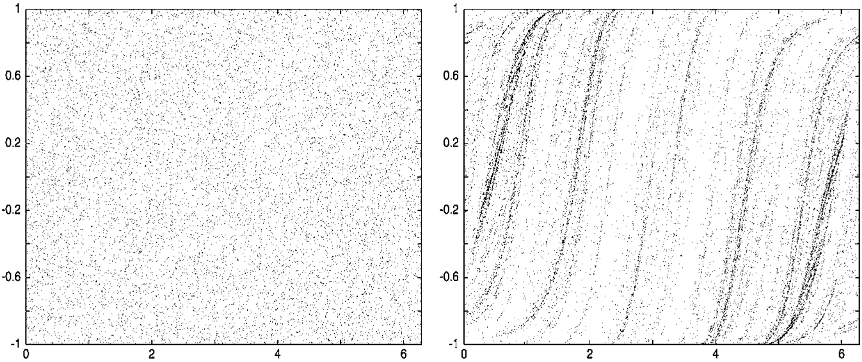

Fig. 1. Distributions of 6×10^6 points (ϕ, η) generated by the dynamics of the periodic nonequilibrium Lorentz gas with $E = 0.1$ (left panel) and $E = 1$ (right panel). Decreasing E, less and less structured distributions are obtained, so that $E = 0$ corresponds to the uniform distribution. However, it is obvious that distributions which cannot be distinguished from the uniform one, in practice, can be obtained with relatively high fields.

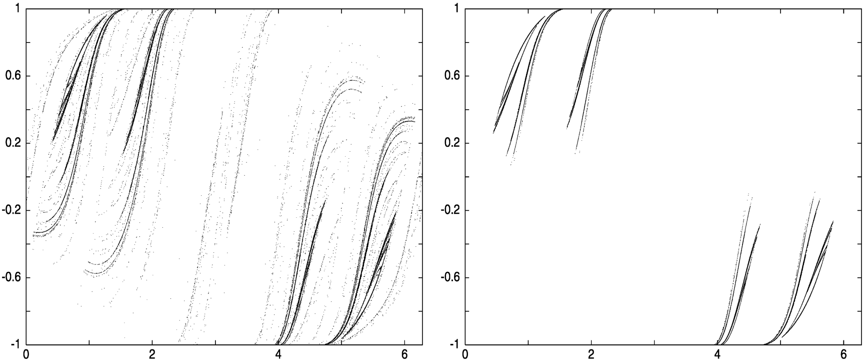

Fig. 2. Distributions of 6×10^6 points (ϕ, η) generated by the dynamics of the periodic nonequilibrium Lorentz gas with $E = 2.1$ (left panel) and $E = 2.2$ (right panel). For fields between $E = 1$ and $E = 2$, the support of the distribution changes gradually, concentrating on smaller and smaller sets. Differently, the transition from $E = 2.1$ to $E = 2.2$ is abrupt, as described in refs. 14 and 15, but these are exceedingly high fields.

Therefore, we argue that there is a level of description within which some quantity related to the Gibbs entropy should change little in the passage from equilibrium to near equilibrium states, similarly to the physical entropy. This poses the problem of the definition of one such quantity that could represent the physical entropy of a particle system in near equilibrium steady states. This is the question investigated in the next section.

5. GREEN'S EXPANSION FOR THE ENTROPY

As remarked above, the dimensional reduction that occurs in driven particle systems modeling real nonequilibrium thermodynamic systems is tiny. Moreover, much more generally than just in the case of the Lorentz gas, small fields should produce only small modifications in the corresponding phase space distributions, despite the abrupt changes suggested by the transition from absolutely continuous to singular distributions. This disparity suggests that if the Gibbs entropy were written as a series involving distribution functions of reduced order, then we might be able to obtain a better understanding of the divergence of the steady state Gibbs entropy, and perhaps even a useful way of renormalizing the divergence (Chapt. 10, ref. 1).

For simplicity we assume that all particles are identical. This means that all particles must feel the combined effects of the dissipative external field and the ergostat. To retain separate wall regions means that the system must be treated as a mixture.

H. S. Green(21) used Kirkwood's factorization (1942) of the N-particle distribution function to write an expansion for the entropy. If

$$f_m(\Gamma_1,\dots,\Gamma_m,t)\equiv\int d\mathbf{q}_N\,d\mathbf{p}_N\cdots d\mathbf{q}_{m+1}\,d\mathbf{p}_{m+1}\,f(\Gamma,t)\qquad m=1,\dots,N \tag{25}$$

are the partial m-body distribution functions, and we define the N z-functions as

$$\begin{aligned}
\ln f_1(\Gamma_1,t)&\equiv z_1(\Gamma_1,t)\\
\ln f_2(\Gamma_1,\Gamma_2,t)&\equiv z_2(\Gamma_1,\Gamma_2,t)+z_1(\Gamma_1,t)+z_1(\Gamma_2,t)\\
\ln f_3(\Gamma_1,\Gamma_2,\Gamma_3,t)&\equiv z_3(\Gamma_1,\Gamma_2,\Gamma_3,t)+z_2(\Gamma_1,\Gamma_2,t)+z_2(\Gamma_1,\Gamma_3,t)\\
&\quad+z_2(\Gamma_2,\Gamma_3,t)+z_1(\Gamma_1,t)+z_1(\Gamma_2,t)+z_1(\Gamma_3,t)\\
&\ \ \vdots
\end{aligned}\tag{26}$$

then Green showed that the Gibbs' entropy can be written as the sum:[7]

$$\begin{aligned}
S_G(t)&=-Nk_{\mathrm{B}}\left\{\frac{1}{1!}\int d\Gamma_1\,f_1(\Gamma_1,t)\,z_1(\Gamma_1,t)\right.\\
&\qquad\left.+\frac{1}{2!}\int d\Gamma_1\,d\Gamma_2\,f_2(\Gamma_1,\Gamma_2,t)\,z_2(\Gamma_1,\Gamma_2,t)+\cdots\right\}\\
&=-Nk_{\mathrm{B}}\left\{\begin{array}{l}
\frac{1}{1!}\int d\Gamma_1\,f_1(\Gamma_1,t)\ln[f_1(\Gamma_1,t)]\\
\qquad+\frac{1}{2!}\int d\Gamma_1\,d\Gamma_2\,f_2(\Gamma_1,\Gamma_2,t)\\
\qquad\times\ln[f_2(\Gamma_1,\Gamma_2,t)/(f_1(\Gamma_1,t)\,f_1(\Gamma_2,t))]\\
\qquad+\cdots
\end{array}\right\}\\
&\equiv(S_{(1)}(t)+S_{(2)}(t)+S_{(3)}(t)+\cdots+S_{(N)}(t)),
\end{aligned}\tag{27}$$

with the shorthand notation $\Gamma_i=(\mathbf{q}_i,\mathbf{p}_i)$. At equilibrium this expansion converges rapidly. In a Lennard-Jones fluid the first 2 terms contribute more than 80% of the total entropy over the entire fluid regions of the equilibrium phase diagram.(22, 23) We note that in a thermodynamic system

---

[7] If we wish to retain distinct thermostatting walls, the Green expansion entails a double sum over wall and system particles. There are no fundamental differences in the general conclusions from both approaches.

with pair interactions (e.g., a Lennard-Jones system), other thermodynamic properties depend only on low order distribution functions. For example, energy and pressure are only dependent on singlet and pair distribution functions, the specific heats and the thermodynamic temperature $(\partial E/\partial S|_V)$ only on singlet, pair and triplet distributions. Equilibrium thermodynamics thus involves only the very lowest order distribution functions defined in an exceedingly high dimensional phase space. Combining these observations with the fact that for Newtonian fluids under typical flow conditions the dimensional decrease of the steady state attractor is tiny (1 part in 10^{23}!), we foresee a reconciliation of the divergence of the Gibbs entropy and the hypothesis of local thermodynamic equilibrium.

In nonequilibrium steady states, the definition (25) of the partial m-body distribution functions cannot be used because they are based on the existence of a full phase space density $f_N(\Gamma, t)$, which does not exist. However, the notion of a partial m-body distribution still makes sense in a system of N identical *but interacting* particles. The partial m-body distribution describes the steady state statistics of the motion of the any m particles $(m \ll N)$, (i.e., as the invariant probability measure μ_m in the m-particle space Ω_m of coordinates $(\Gamma_1,..., \Gamma_m)$). It seems clear that for $m \ll N$ the μ_m must have a density f_m, since they are obtained projecting down the full phase space probability distribution onto subspaces of much lower dimension, hence performing many regularizing operations. In other words, the low order probability distributions lose information about the finely detailed, possibly fractal, full phase space distribution, and should not be singular.[8]

This hypothesis cannot be checked in many particle systems, because of the difficulty of numerically constructing probability distributions in high dimensional spaces. However, using the Poincaré map considered in the previous section, we can support the idea that projections produce regular distributions. Let $\mu_E(d(\phi, \eta)$ be the stroboscopic phase space distribution, which is known to be singular for non-vanishing E, and to be uniform for $E = 0$. We construct the two distributions $\lambda_E(d\phi)$ and $\nu_E(d\eta)$, building the histograms of the R values taken by the variables ϕ and η on a simulation of R collisions, using partitions of the intervals I_ϕ and I_η made of B cells of equal size. Then, for fixed B, we let R increase to obtain the probability distributions corresponding to the partitions made of B cells, and successively we let B grow to approximate λ_E and ν_E better and better.

---

[8] In this regularisation we assume that there is no connection between the singularity structures apparent in f_N, and the particle index. Such connections *do* exist in the *noninteracting* N-particle Lorentz gas. For this system the very obvious connection is, $\mu_N = \bigotimes_{i=1}^{N} \mu_1$. However, for interacting N-particle systems, no such connections are expected.

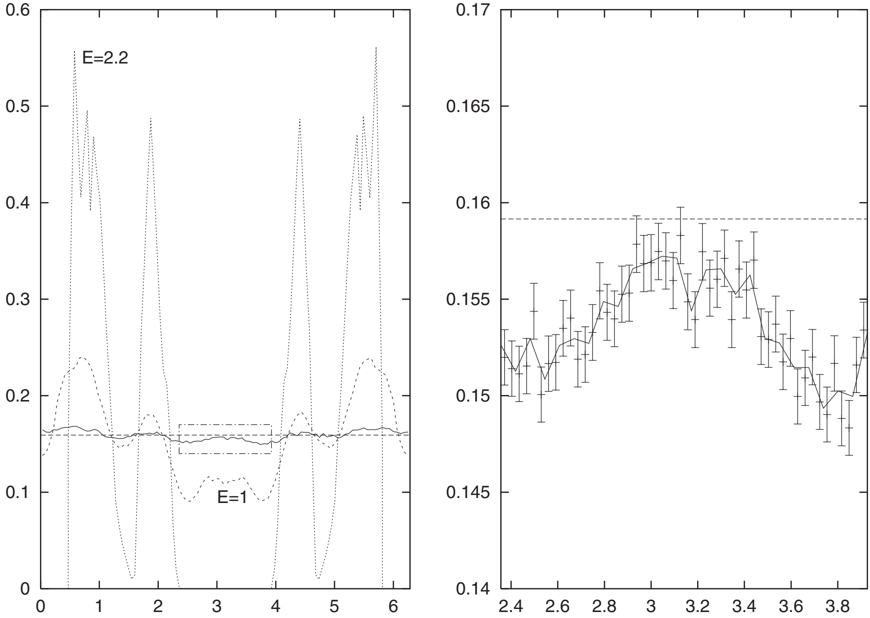

Fig. 3. Distributions of the variable ϕ for $E = 0$, 0.1, 1, and 2.2, obtained with $B = 100$ subdivisions of the interval I_ϕ (left panel). The straight horizontal line represents the uniform equilibrium distribution. The number of points used is $R(E = 0.1) = 2.344 \times 10^6$, $R(E = 1) = 3.144 \times 10^6$, and $R(E = 2.2) = 2.745 \times 10^6$. The right panel is a magnification of the rectangular box in the left panel containing a piece of the distributions with $E = 0, 0.1$. Here, the piecewise linear curve is the distribution for $E = 0.1$ portrayed in the left panel, while the points with error bars represent the corresponding distribution obtained with $B = 200$ and same R. For the sake of clarity, only the two smallest values of B used have been reported, but the same features have been observed up to B of order $O(10^4)$.

The results of this procedure are portrayed in Figs. 3 and 4 for $E = 0.1$, 1, and 2.2. Figure 3 shows the histograms of the variable ϕ, while Fig. 4 those of the variable η and, in both, the horizontal straight line represents the equilibrium uniform distribution. The histograms are not particularly smooth, but consistent with the hypothesis that λ_E and ν_E do have a density, which we denote by $P_{E,\phi}$ and $P_{E,\eta}$ respectively. For instance, the right panels of these figures show that, if we fix R and let B grow, the histograms become more irregular; but this happens because the same number of events must be used in a larger number of bins, giving rise to less accurate distributions. This loss of accuracy is reflected on the error bars which grow, keeping within themselves any distribution obtained with

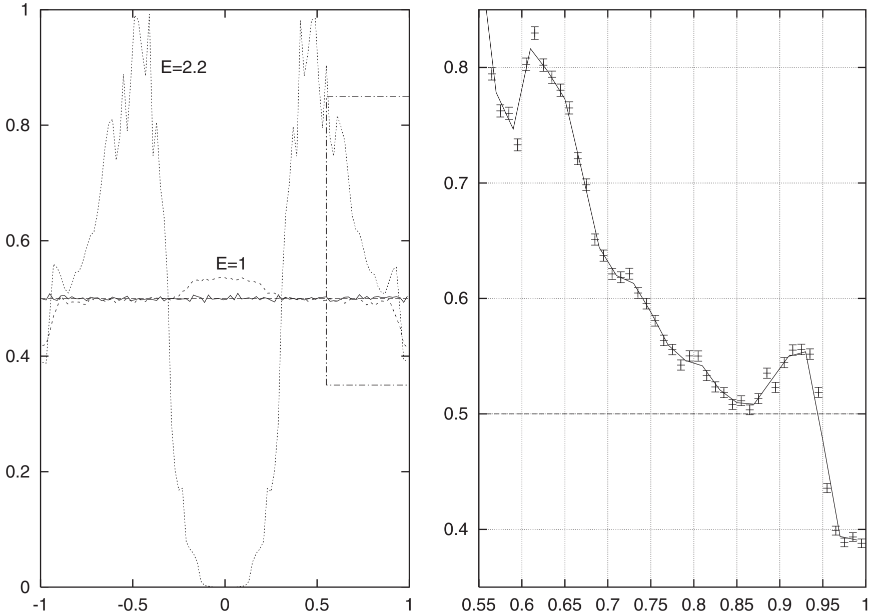

Fig. 4. Distributions of the variable η for $E = 0$, 0.1, 1, and 2.2, obtained with $B = 100$ subdivsions of the interval I_η (left panel). The straight horizontal line represents the uniform equilibrium distribution. The number of points used to build these distributions is the same as in Fig. 3. The right panel is a magnification of the rectangular box in the left panel containing a piece of the distributions only for the part $E = 0$ and $E = 2.2$. Here, the piecewise linear curve is the distribution for $E = 2.2$ portrayed in the left panel, while the points with error bars represent the corresponding distribution obtained with $B = 200$ and same R. Similarly to Fig. 3, only the two smallest values of B used have been reported.

sufficiently large B and R.[9] If, on the other hand, we fix B and let R grow, no sensible variation on the distribution is noticed after R has exceeded a certain value R_B, which depends on B. As a consequence, the quantities

$$S_\phi(E) = -\int_{I_\phi} d\phi' \, P_{E,\phi}(\phi') \ln P_{E,\phi}(\phi'); \quad S_\eta(E) = -\int_{I_\eta} d\eta' \, P_{E,\eta}(\eta') \ln P_{E,\eta}(\eta') \tag{28}$$

are expected to take finite values as, indeed, seems to be the case. We approximated these integrals using the histograms mentioned above, and then we varied both B and R, to see if the corresponding approximate

---

[9] Note that we have drawn only one set of error bars, for the sake of figure readability. Had we drawn both sets of error bars in the right panels of Figs. 3 and 4, the overlap of the distributions obtained with different B would have looked even more complete.

values converge to definite values for each E. For every field E and fixed simulation length R, we found that S_ϕ and S_η decrease as B grows, but also that the rate of decrease is slower for larger R, suggesting that such ratio would vanish in the $R \to \infty$ limit. Indeed, at least for not too large E, S_ϕ and S_η are easily seen to converge to given *positive* values when B is kept fixed and R is increased. In particular, for $E = 0.1$ these limiting values are $S_\phi(0.1) = 1.8374(1)$ and $S_\eta(0.1) = 0.6931(1)$ (i.e., with an uncertainty of ± 1 in the fourth digit), which are very close to the equilibrium values $S_\phi(E = 0) = 1.83788$ and $S_\eta(E = 0) = 0.693147$, obtained from the projected uniform distributions. For the largest E's, it is not equally easy to extrapolate S_ϕ and S_η in the limit $R \to \infty$, but it looks quite plausible that one such limit exists and is a positive number, as evidenced also by Fig. 5. There, each panel gives S_ϕ as a function of B for several values of R: the left panel refers to $E = 0.1$ while the right panel refers to $E = 2.2$! The

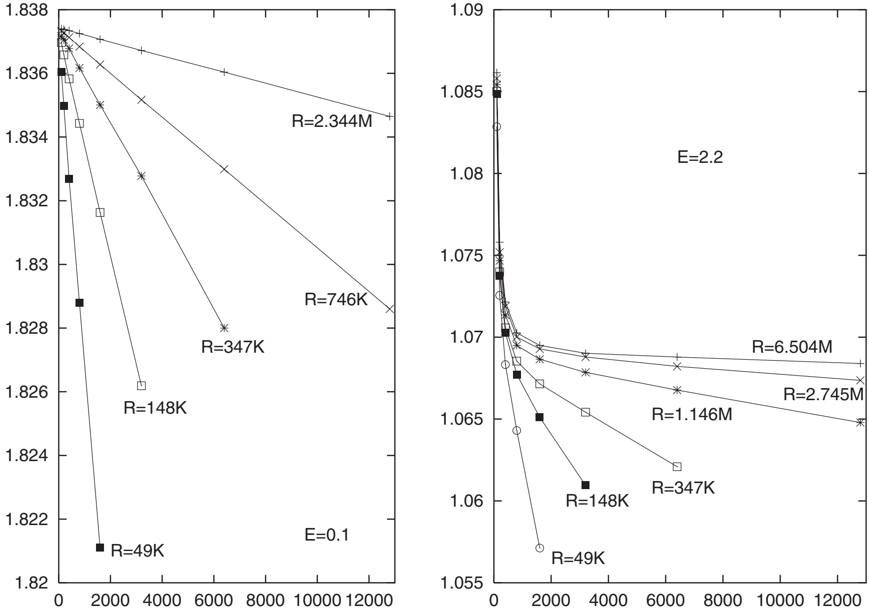

Fig. 5. Values of S_ϕ as a function of B for $E = 0.1$ (left panel) and $E = 2.2$ (right panel). The different curves are obtained with different values of R, as indicated by the labels, in which K stands for 10^3 and M for 10^6. The exact values of S_ϕ are obtained taking first the $R \to \infty$ limit and later the $B \to \infty$ limit. This procedure yields $S_\phi(E = 0.1) = 1.8374(1)$ with a small numerical effort, while $S_\phi(E = 2.2)$ needs much longer runs, presently out of our reach. However, it is quite reasonable to conclude that also in the case $E = 2.2$ the limit S_ϕ exists and is a positive number.

results for intermediate E's are similar, and appear to connect smoothly the data of these two extreme cases.

Therefore, in accord with our conjectures, the projected distributions appear to be non-singular even at rather large E, and to yield values S_ϕ and S_η close to the equilibrium ones if E is small, despite the fact that the stroboscopic phase space distribution is singular, and the corresponding Gibbs entropies not defined for *any* $E > 0$.

On the grounds of these observations, it is of interest to ask, for many particle systems, what is the highest dimensionality of the projection with nonsingular projected distribution? We propose the following:

Definition. For dissipative N-particle systems, let the Green dimension, N_G, be $2d_C$ times the largest integer for which the corresponding partial contribution to Green expansion for the entropy remains finite, i.e.,

$$\lim_{t\to\infty} |S_{(m)}(t)| = |S^{\infty}_{(m)}| < \infty, \qquad \text{for} \quad 2d_C m \leqslant N_G < 2d_C N \tag{29}$$

$$\lim_{t\to\infty} |S_{(m)}(t)| = \infty, \qquad \text{for} \quad N_G < 2d_C m \leqslant 2d_C N. \tag{30}$$

Our discussion of the Kaplan–Yorke dimension suggests that N_G should be approximately equal to N_{KY}. However, as noted earlier, one obvious violation of this idea is afforded by the systems of independent (non-interacting) particles, including the many particle Lorentz gas made of many replicas of the same 1-particle system discussed in ref. 17. Indeed, the phase space distributions of such systems factorize as products of 1-particle distributions, each of which is singular in its $3d$-dimensional phase space. Therefore, the integration over any number m of particle coordinates does not remove the singular character of the corresponding reduced distributions and $S_{(m)}$ is not defined. Although non-interacting particle systems may be pedagogically useful, they are in fact, ensembles of 1-particle systems, as pointed out in ref. 17 and not even as idealized models, relevant in the description of near equilibrium thermodynamics. In fact, due to the lack of collisions, non-interacting particle systems do not give rise to local equilibrium states, which is a prerequisite for the thermodynamic quantities to be defined.(14) [10]

Hence, we conjecture that for interacting particle systems the Green dimension exists and equals the Kaplan–Yorke dimension, i.e., the

---

[10] In the case of the Lorentz gas, the elastic collisions of the particles with the fixed scatterers cannot give rise to local Maxwellian distributions or to any thermodynamic force.

discriminating dimension for which (29), (30) hold is $N_G = N_{\mathrm{KY}}$. At small fields, where linear irreversible thermodynamics applies, we expect that

$$\lim_{t \to \infty} S_{(m)}(t; F_e) = S^{\infty}_{(m)}(F_e) \approx S^{\infty}_{(m)}(F_e = 0) + O(F_e^2), \qquad \forall m < N. \tag{31}$$

This can be reconciled with the equation of change for the Gibbs entropy only if

$$\lim_{F_e \to 0} \lim_{t \to \infty} S_N(t; F_e) = \lim_{F_e \to 0} S^{\infty}_N(F_e) = -\infty \tag{32}$$

and

$$\lim_{F_e \to 0} \lim_{t \to \infty} \dot{S}_{(N)}(t; F_e) = \lim_{F_e \to 0} \lim_{t \to \infty} \dot{S}_G(t; F_e) = -\lim_{F_e \to 0} \lim_{t \to \infty} dN\, k_{\mathrm{B}} \langle \alpha(t) \rangle = 0. \tag{33}$$

In other words we argue that in the zero field limit, where linear irreversible thermodynamics and Green–Kubo relations are expected to be valid, the entire divergence of the Gibbs entropy is carried by the last term in the Green-expansion. All the terms that are characteristic of low order distribution functions are finite and differ from their equilibrium values by time independent amounts that are proportional to F_e^2. For sufficiently small fields these differences in the thermodynamic entropy are insignificant and the system is said to be in local thermodynamic equilibrium. At the same time the Gibbs entropy diverges in the long time limit, towards negative infinity. We argue that this divergence cannot be seen in the thermodynamic entropy of the system because it is manifest *only* in extremely high order distribution functions. This divergence does however, liberate entropy to the surroundings: the so-called spontaneous entropy production of linear irreversible thermodynamics. For constant internal energy systems in a steady state, this entropy production is equal in magnitude but opposite in sign, to the rate of change of the system's Gibbs entropy. Therefore, the—to many extents uninteresting—singularity of the phase space distribution acquires significance determining the steady state rate of energy dissipation into heat, for the NEMD models whose phase space contraction rate can be identified with the entropy production rate.

6. CONCLUSION

The definition of entropy away from equilibrium has been a central problem in statistical mechanics since the time of Gibbs himself. The most common "solution" to this problem has been to attempt to define a coarse grained Gibbs entropy (pioneered by Gibbs,$^{(10)}$ see also refs. 24 and 25 but note the cautions given in ref. 26).

Our approach [ref. 1, Chapt. 10] is quite different. As stressed in the literature (see, e.g., refs. 1, 26, and 27), the reason for the divergence of the Gibbs entropy is the loss of dimensionality in the steady state attractor. Using the fact that in macroscopic thermodynamic systems this dimension loss is miniscule, we use H. S. Green's 1952 expansion of the Gibbs entropy in terms of partial distribution functions to approximate the thermodynamic entropy as a series of integrals over smooth lower dimensional functions. Until the dimension of these integrals approaches the dimension of the steady state attractor ($\sim$ Kaplan–Yorke dimension), each integral is expected to be well behaved. This procedure is extraordinarily difficult to carry out in practice,$^{(1, 22, 23)}$ but it is expected to lead to a well defined *reduced* Gibbs entropy because the entropy is estimated from *within* the dimensionality of the steady state attractor.

In the zero field limit where the dimension loss is quadratic in the dissipative field, our approach leads to a straightforward understanding of Maxwell's postulate of local thermodynamic equilibrium. The reduced Gibbs entropy converges to the equilibrium Gibbs entropy quadratically in the field. In this limit the entire divergence of the Gibbs nonequilibrium entropy is carried by the last term in the Green entropy expansion which involves distribution functions which are of the order of Avogadro's number!

The entropy absorbed by the thermostat is seen to result from the collapse of the distribution function for the entire system towards a steady state attractor that has lower dimension than the ostensible dimension of phase space. This "spontaneous entropy production" is seen to be the most important physical consequence of the loss of dimensionality.

One might object that our analysis is invalidated by our use of an artificial (time reversible) ergostat. However this thermostat can be made arbitrarily remote from the system of physical interest.$^{(8)}$ This system cannot "know" the precise details of how entropy was removed at such a remote distance. This means that the results obtained for the system using our simple mathematical ergostat should not differ substantially from those we would infer for the same system surrounded (at a distance) by a real physical thermostat (say with a huge heat capacity), in accord with emerging theory on the equivalence of nonequilibrium ensembles mentioned in the introduction.

One should note that the form of our equations of motion (2) is quite general. In particular this form encompasses the possibility of treating ergostatted Rayleigh–Benard flows or other spatially inhomogeneous systems. In these spatially inhomogeneous systems, (6) still holds—the entropy absorbed (emitted) by the thermostatting walls and defined in (5), is still related to the negative of the rate of change of the Gibbs entropy for the entire system. Further if such spatially inhomogeneous systems do have a steady state then we expect that the relationship of the rate of change of the Gibbs entropy to the Lyapunov spectra and the Kaplan–Yorke and Green dimensions still holds.

For wall thermostatted rather than homogeneously thermostatted systems the mathematical treatment of the Green expansion will become much more complex with the entropy being written as a double sum over system and wall particles. However, we still expect that when the dimension of the integrals in a term appearing in the expansion is less than the Kaplan–Yorke dimension, that term will make a finite contribution to the sum. Conversely if the dimension of the relevant integral is greater than the Kaplan–Yorke dimension that term is expected to be infinite. We remark that calculating the entropy production in an inhomogeneous system by spatially integrating the local entropy source strength is a complex task. Also, in such systems although local kinetic temperature inside the walls may vary from place to place, nevertheless (5) still holds.

One might ask how long it takes the accessed phase space to collapse onto the steady state attractor. From (5), (20) we see that in the small field limit

$$V_\Gamma(t) \approx V_\Gamma(0) \exp[-L(F_e = 0)\, V F_e^2 t / k_B T] \tag{34}$$

where $L(F_e = 0)$ is the limiting zero field transport coefficient. This would suggest that in the zero field limit the characteristic time for the collapse onto the steady state attractor diverges to infinity as F_e^2. However this is misleading because as the field goes to zero the dimension of the collapsing subspace also goes to zero. The ostensible phase space volume element, V_Γ, is the product of an invariant Kaplan–Yorke volume whose dimension is of course D_{KY}, and a collapsing volume whose dimension is $2d_C N - D_{\mathrm{KY}} = O(F_e^2)$. Asking how fast V_Γ collapses onto the attractor is equivalent to asking how rapidly the collapsing volume, collapses. Canceling the invariant Kaplan–Yorke volume from both sides of (34) and using (23), we can write the evolution of the collapsing volume as

$$L_{\mathrm{col}}^{(2d_C N - D_{\mathrm{KY}})}(t) \approx L_{\mathrm{col}}^{(2d_C N - D_{\mathrm{KY}})}(0) \exp[-\lambda_{\max}(F_e = 0)(2d_C N - D_{\mathrm{KY}}(F_e))\, t], \tag{35}$$

in the small field, long time limit. Now it is trivial to see that in this same limit, the equation of change for the collapsing length is just

$$L_{\mathrm{col}}(t) \approx L_{\mathrm{col}}(0) \exp[-\lambda_{\max}(F_e = 0)\, t]. \tag{36}$$

This leads us to conjecture the following: the length associated with the limiting small field collapse towards the steady state attractor decreases at a rate that is independent of the magnitude of the external field and the ostensible dimension of phase space. This rate is in fact as fast as it can possibly be since it is given by the *largest* Lyapunov exponent.

The fundamental significance of entropy in statistical thermodynamics results from the fact that this quantity satisfies a number of extremum and variational properties both at equilibrium and close to equilibrium where linear irreversible thermodynamics is valid.(14) We now recall that even far from equilibrium, the entropy absorbed by the ergostat Σ—Eq. (5), satisfies fundamental and quite general Fluctuation Theorems. Take for example, an ensemble of isoenergetic systems that is initially at equilibrium and subject to ergostatted dissipative dynamics for $t > 0$, (2). It has been proved(9, 28–30) that the time averaged entropy absorption $\bar{\Sigma}(t)$ satisfies fluctuation relations which show that as the system size and the averaging time t increase it becomes exponentially more likely that the entropy absorbed by the ergostat is positive rather than negative.

ACKNOWLEDGMENTS

One of us (D.J.E.) would like to thank the staff at NIST, Boulder, Colorado, for their support of much of this work. L.R. acknowledges financial support from INFM and from GNFM-INDAM. We would both like to thank Professor J. R. Dorfman for inspiration and support over many years. Thanks for your help Bob!

REFERENCES

1. D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Liquids* (Academic Press, New York 1990); also, http://rsc.anu.edu.au/~evans/evansmorrissbook.htm.
2. S. Sarman, D. J. Evans, and P. T. Cummings, *Physics Reports*, M. J. Klein, ed. (Elsevier, 1998), **305**(1–2):1.
3. G. Gallavotti, *Physica D* **105**:163 (1997).
4. J. Vollmer, T. Tel, and W. Breymann, *Phys. Rev. Lett.* **79**:2759 (1997).
5. L. Rondoni and E. Segre, *Nonlinearity* **12**:1471 (1999).
6. G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Experimental demonstration of violations of the second law of thermodynamics, *Phys. Rev. Lett.* (submitted).
7. S. Ciliberto and C. Laroche, *J. Physique IV* **8**(6):215 (1998).
8. G. Ayton, D. J. Evans, and D. J. Searles, *J. Chem. Phys.* **115**:2033 (2001).
9. D. J. Evans and D. J. Searles, *Phys. Rev. E* **50**:1645 (1994).
10. J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (Yale University Press, Boston, 1902).

11. B. D. Butler, G. Ayton, O. G. Jepps, and D. J. Evans, *J. Chem. Phys.* **109**:6519–6522 (1998).
12. J. L. Kaplan and J. A. Yorke, *Comm. Math. Phys.* **67**:93 (1979).
13. D. J. Evans, E. G. D. Cohen, D. J. Searles, and F. Bonetto, *J. Stat. Phys.* **101**:17 (2000).
14. S. R. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (Dover Pub., New York, 1984).
15. B. Moran and W. G. Hoover, *J. Stat. Phys.* **48**:709 (1987).
16. J. Lloyd *et al.*, *Chaos* **5**(3):536 (1995).
17. E. G. D. Cohen and L. Rondoni, *Chaos* **8**(2):357 (1998).
18. C. P. Dettmann and G. P. Morriss, *Phys. Rev. E* **54**:4782(1996).
19. Ch. Dellago *et al.*, *Phys. Rev. E* **52**(5):4817 (1995).
20. G. P. Morriss, C. P. Dettmann, and L. Rondoni, *Physica A* 240:84 (1997).
21. H. S. Green, *The Molecular Theory of Fluids* (North-Holland, Amsterdam, 1952); see also, H. J. Raveché, *J. Chem. Phys.* **55**:2242 (1971).
22. A. Baranyai and D. J. Evans, *Phys. Rev. A* **40**:3817 (1989).
23. A. Baranyai and D. J. Evans, *Phys. Rev. A* **42**:849 (1990).
24. P. Gaspard, *J. Stat. Phys.* **88**:1215 (1997). J. Vollmer, T. Tell, and W. Breymann, *Phys. Rev. E* **58**:1672 (1998). J. Vollmer and W. Breymann, *Phys. Rev. Lett.* **79**:2759 (1997). W. Breymann and J. Vollmer, *Chaos* **8**:396 (1998).
25. T. Gilbert and J. R. Dorfman, *J. Stat. Phys.* **96**:225 (1999).
26. L. Rondoni and E. G. D. Cohen, *Nonlinearity* **13**:1905 (2000).
27. S. Goldstein, J. L. Lebowitz, and Ya. Sinai, *Chaos* **8**(2):393 (1998).
28. D. J. Evans, E. G. D. Cohen, and G. P. Morriss, *Phys. Rev. Lett.* **71**:2401–2404 (1993); Erratum, **71**:3616 (1993).
29. G. Gallavotti and E. G. D. Cohen, *Phys. Rev. Lett.* **74**:2694 (1995).
30. D. J. Searles and D. J. Evans, *J. Chem. Phys.* **113**:3503–3509 (2000).