Ergodicity of non-Hamiltonian Equilibrium Systems

Denis J. Evans[1,2], **Stephen R. Williams**[1,3], **Lamberto Rondoni**[4,5,6], **Debra J. Searles**[1,6,7]

[1] *Australian Institute for Bioengineering & Nanotechnology*
The University of Queensland, Brisbane QLD, 4072, Australia

[2] *Department of Applied Mathematics, Research School of Physics and Engineering*
Australian National University, Canberra ACT, 0200, Australia

[3] *Research School of Chemistry, Australian National University*
Canberra ACT, 0200 Australia

[4] *Dipartimento di Scienze Matematiche and Graphene@Polito Lab, Politecnico di Torino*
Corso Duca degli Abruzzi 24, 10129 Torino, Italy

[5] *INFN, Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy*

[6] *Kavli Institute for Theoretical Physics, Chinese Academy of Sciences, China*

[7] *School of Chemistry and Molecular Biosciences*
The University of Queensland, Brisbane QLD, 4072, Australia

Abstract: It is well known that ergodic theory can be used to formally prove a form of relaxation to microcanonical equilibrium for finite, mixing Hamiltonian systems. In this manuscript we substantially modify this proof using an approach similar to that used in umbrella sampling, and use this approach to consider relaxation in both Hamiltonian and non-Hamiltonian systems. In doing so, we demonstrate the need for a form of ergodic consistency of the initial and final distribution. The approach only applies to relaxation of averages of physical properties and low order probability distribution functions. It does not provide any information about whether the full $6N$-dimensional phase space distribution relaxes towards the equilibrium distribution or how long the relaxation of physical averages takes.

Key words: equilibrium, ergodic theory, relaxation to equilibrium, distribution function, statistical mechanics

I. INTRODUCTION

We have recently presented a proof of relaxation to microcanonical equilibrium for a Hamiltonian system utilizing the ergodic theorem and an approach based on umbrella sampling [1]. This approach can be extended to apply to other ensembles, as we discussed in reference [2]. In this paper we reproduce these arguments, and address the paradox that although the properties of a system relax to their equilibrium values under appropriate conditions, the distribution function never evolves to the $6N$-dimensional equilibrium distribution function. This is resolved by considering lower-order distributions.

It is known from ergodic theory that a finite, autonomous, Hamiltonian system that preserves a mixing microcanonical equilibrium distribution, will, from almost any initial distribution of states described by a phase space distribution, eventually relax towards a situation where averages

---

[1] In systems without boundaries the initial distribution is $f(\mathbf{\Gamma}; 0)\delta(H(\mathbf{\Gamma}) - E)\delta(\mathbf{P})$, where $\mathbf{P}$ is the total linear momentum, $H(\mathbf{\Gamma})$ is the Hamiltonian and E the energy of the system.

of physical phase functions approach their microcanonical equilibrium values [3].[1] These statements refer to an *equilibrium* microcanonical *distribution* that has been defined in [4,5], and is discussed further in Appendix A.

A system is said to be *mixing* if for integrable, phase functions, time correlation functions computed with respect to a stationary distribution factorize into products of averages computed with respect to the same distribution [3]:

$$\lim_{t\to\infty} \left\langle A(\mathbf{\Gamma})B(S^t\mathbf{\Gamma})\right\rangle_\infty - \left\langle A(\mathbf{\Gamma})\right\rangle_\infty \left\langle B(\mathbf{\Gamma})\right\rangle_\infty = 0. \quad (1)$$

Here the brackets, $\langle ... \rangle_\infty$, denote an ensemble average with respect to an invariant (*i.e.* time-stationary) probability measure μ_∞ and $S^t\mathbf{\Gamma}$ is the position of the evolved phase vector $\mathbf{\Gamma}$ after time t. In the case that μ_∞ has a density , one may write:

$$\langle A\rangle_\infty = \int d\mu_\infty(\mathbf{\Gamma})A(\mathbf{\Gamma}) = \int d\mathbf{\Gamma}\, f(\mathbf{\Gamma};\infty)A(\mathbf{\Gamma}), \quad (2)$$

where $f(\mathbf{\Gamma};\infty) = \partial\mu_\infty/\partial\mathbf{\Gamma}$ is a normalized distribution and $f(\mathbf{\Gamma};\infty)d\mathbf{\Gamma}$ is dimensionless.[2]

Implicit in this definition is the fact that the invariant measure must be preserved by the dynamics, in which case at all times $\langle B(S^t\mathbf{\Gamma})\rangle_\infty = \langle B(\mathbf{\Gamma})\rangle_\infty$ for all phase variables, $B(\mathbf{\Gamma})$. So the notion of mixing systems requires an invariant measure, *i.e.* a measure that is preserved by the dynamics *and* additionally they must satisfy (1) with respect to this invariant distribution or measure.[3] This invariant *measure* implies that the averages of physical quantities are invariant.

The mixing property is a property of the stationary state of interest, in which observables take the average values denoted by $\langle ... \rangle_\infty$. It represents the fact that, in the macroscopically stationary state, correlations among microscopically evolving physical properties (physically relevant phase functions) decay in time. Therefore, in general the mixing condition *does not* guarantee relaxation to an invariant state starting from a macroscopic non-invariant state. Mixing already *assumes* stationarity of the macrostate and its preservation by the system's dynamics, regardless of whether it is reached asymptotically in time, as implied by our notation, or it is initially prepared in that state by some means (*e.g.* a Monte Carlo process).

II. ERGODIC THEORY PROOF OF RELAXATION FOR AUTONOMOUS HAMILTONIAN SYSTEMS

For completeness we repeat here our version [1] of the standard ergodic theory proof of relaxation for autonomous Hamiltonian systems. We begin by noting that the *microcanonical distribution*, $f_{\mu c}(\mathbf{\Gamma})$ for the infinitesimally thin energy shell, $D(N,V,E,\delta E): E < H(\mathbf{\Gamma}) < E + \delta E$ which we refer to as D, is given by:

$$f_{\mu c}(\mathbf{\Gamma}) \equiv \begin{cases} \dfrac{1}{\int_D d\mathbf{\Gamma}}, & \mathbf{\Gamma} \in D \\[2mm] 0, & \mathbf{\Gamma} \notin D \end{cases} \quad (3)$$

The microcanonical distribution function, $f_{\mu c}(\mathbf{\Gamma})$, is a time-stationary distribution, preserved by the autonomous Hamiltonian dynamics [5]. One can see from the phase continuity equation (often referred to as the Liouville equation) that both $df_{\mu c}(\mathbf{\Gamma})/dt = 0$ and $\partial f_{\mu c}(\mathbf{\Gamma})/\partial t = 0$.[4] For a discussion on the need to consider a thin shell rather than a constant energy hypersurface, see Appendix B.

We will now give the standard proof that if our ensemble is initially *not* distributed according to this distribution, the ensemble will in fact relax towards this distribution – at least as can be ascertained by computing time averages of low order physical phase functions.[5]

The time dependent average of a physical phase function $A(\mathbf{\Gamma})$ for some smooth distribution function, $f(\mathbf{\Gamma};t)$ that evolves in time according to the Hamiltonian dynamics and which is defined in the domain, *D,* is given by:

$$\begin{aligned}\langle A\rangle_t &= \int_D d\mathbf{\Gamma} A(\mathbf{\Gamma})f(\mathbf{\Gamma};t) \\ &= \int_D d\mathbf{\Gamma} A(S^t\mathbf{\Gamma})f(\mathbf{\Gamma};0),\end{aligned} \quad (4)$$

where the second equality is due to the equivalence of the Heisenberg and Schrödinger representations of phase space averages [6] and the notation $\langle A\rangle_t$ refers to an ensemble average with respect to the time evolved distribution . In (4) the stationarity of the distribution $f(\mathbf{\Gamma};t)$ is not assumed (so $f(\mathbf{\Gamma};t)$ is not necessarily equal to $f(\mathbf{\Gamma};0)$). However, since the dynamics is driven by an autonomous Hamiltonian, the energy is fixed. Formally, calculation of averages in (4) requires either a knowledge of the full, time dependent, phase

---

[2] If on the other hand, if μ_∞ is singular with respect to the Lebesgue measure there is no density and one can only write the first equality, $\langle A\rangle_\infty = \int d\mu_\infty(\mathbf{\Gamma})A(\mathbf{\Gamma})$, where $d\mu_\infty(\mathbf{\Gamma})$ is dimensionless and normalized.

[3] We note that if the system has nonzero angular momentum, no stationary state is possible (unless we transform to a non-inertial, co-rotating coordinate frame where Hamiltonian dynamics breaks down). So if angular momentum is conserved in our system we must set it to zero as can be done for the total linear momentum.

[4] We note that if the dynamics has constant (zero) total momentum one of the particles may be removed from the phase space since its position and momentum is known from the $3(N-1)$ coordinates and momenta in the system. This also removes a source of non-independence in the phase space variables that would otherwise make the evaluation of $(\partial/\partial\mathbf{\Gamma}) \bullet \dot{\mathbf{\Gamma}}(\mathbf{\Gamma}) = 0$ difficult.

[5] Note: low order physical phase functions can be evaluated using functionals of low order (singlet, pair and triplet, for example) distribution functions – see Appendix D for details. The full N-particle phase space distribution is a function of ~6($N-1$) variables where N may be of the order of Avogadro's Number, if the system is macroscopic.

space distribution, $f(\boldsymbol{\Gamma}; t)$ or the initial phase space distribution.

Multiplying and dividing the last expression in (4) by the (necessarily finite!) ostensible volume of the phase space casts the first line in a form to which the mixing property can (formally) be applied:

$$
\begin{aligned}
\langle A \rangle_t &= \frac{1}{\int_D d\boldsymbol{\Gamma}} \bullet \int_D d\boldsymbol{\Gamma} \, A(S^t\boldsymbol{\Gamma}) f(\boldsymbol{\Gamma}; 0) \int_D d\boldsymbol{\Gamma} \\
&= \langle A(S^t\boldsymbol{\Gamma}) f(\boldsymbol{\Gamma}; 0) \rangle_{\mu c} \int_D d\boldsymbol{\Gamma}.
\end{aligned}
\tag{5}
$$

We emphasise that in order to derive (5), the value of $\int_D d\boldsymbol{\Gamma}$ needs to be finite and non-zero. We exclude the case where the initial distribution is a nonequilibrium steady state distribution. This is because the measures of these distributions are not smooth.[6]

Knowing that the microcanonical distribution is preserved by the autonomous Hamiltonian dynamics, we now take the long time limit and use the mixing assumption, (1), to allow us to factorize the naturally invariant (microcanonical) time correlation function $\langle A(S^t\boldsymbol{\Gamma}) f(\boldsymbol{\Gamma}; 0) \rangle_{\mu c}$ into a sum of products of two invariant (microcanonical) averages.[7] Then:

$$
\begin{aligned}
\lim_{t \to \infty} \langle A \rangle_t &= \lim_{t \to \infty} \left\langle A(S^t\boldsymbol{\Gamma}) f(\boldsymbol{\Gamma}; 0) \right\rangle_{\mu c} \int_D d\boldsymbol{\Gamma} \\
&= \langle A(\boldsymbol{\Gamma}) \rangle_{\mu c} \langle f(\boldsymbol{\Gamma}; 0) \rangle_{\mu c} \int_D d\boldsymbol{\Gamma} \\
&= \langle A(\boldsymbol{\Gamma}) \rangle_{\mu c} \left[\frac{\int_D d\boldsymbol{\Gamma} f(\boldsymbol{\Gamma}; 0)}{\int_D d\boldsymbol{\Gamma}} \right] \int_D d\boldsymbol{\Gamma} \\
&= \langle A(\boldsymbol{\Gamma}) \rangle_{\mu c}
\end{aligned}
\tag{6}
$$

We note that equation (1) only applies to a stationary state, however we know that a stationary state exists since, as noted above, the microcanonical distribution is preserved by autonomous Hamiltonian dynamics, so this is not an *assumption*. We only have to assume that this invariant distribution is mixing.

The analysis above proves that for mixing systems $\langle A \rangle_t$ tends towards a microcanonical average, whatever physical phase function $A(\boldsymbol{\Gamma})$, or initial probability density $f(\boldsymbol{\Gamma}; 0)$ one considers – as long as it lies in the phase space domain D. This amounts to a formal proof of relaxation *towards* the microcanonical equilibrium state denoted by $\langle \ldots \rangle_{\mu c}$. The proof only shows the relaxation of averages of physical quantities towards their equilibrium values.

The proof given in (4,5,6) shows (formally) that *averages* of thermodynamic quantities *approach* microcanonical averages in the long time limit if the initial distribution is not microcanonical. If the system starts at $t = 0$ in the microcanonical distribution, it becomes a trivial statement since the averages will be the microcanonical averages at all times.

Because distribution functions are normalized we see that for any distribution and time,

$$
\begin{aligned}
\langle f(\boldsymbol{\Gamma}; t) \rangle_{\mu c} &= \lim_{\delta E \to 0} \int_{D(\delta E)} d\boldsymbol{\Gamma} \, f(\boldsymbol{\Gamma}; t) \Big/ \int_{D(\delta E)} d\boldsymbol{\Gamma} \\
&= \lim_{\delta E \to 0} 1 \Big/ \int_{D(\delta E)} d\boldsymbol{\Gamma} = \exp\left[-S_{\mu c}(E, N, V)/k_B\right]
\end{aligned}
\tag{7}
$$

where we use the Gibbs definition of entropy $S_{\mu c} \equiv -k_B \int_D d\boldsymbol{\Gamma} \, f_{\mu c}(\boldsymbol{\Gamma}) \ln[f_{\mu c}(\boldsymbol{\Gamma})]$ and $f_{\mu c}(\boldsymbol{\Gamma}) = \lim_{\delta E \to 0} 1 / \int_{\boldsymbol{\Gamma} \in D(\delta E)} d\boldsymbol{\Gamma}$. Hence, the microcanonical average of a normalized distribution function tells us *nothing* about that distribution. The average only tells us the Gibbs entropy of the microcanonical distribution used to calculate the average.

We know that unlike the averages of phase variables, the phase space density does not relax. For example, if $f(\boldsymbol{\Gamma}_0; 0) = c \neq 0$, then using the streaming version of the phase space continuity equation $f(S^t\boldsymbol{\Gamma}_0; t) = c, \; \forall t > 0$, which would be inconsistent with the eventual relaxation to the microcanonical distribution function at all points in phase space. Equation (6), which says that phase *averages* relax, might seem counter-intuitive because the distribution function never relaxes to the microcanonical distribution. In Section 4 we will treat cases where the initial or final distribution are not fully known, showing how this relates to the relaxation of the distribution function and reconciliation of how phase averages can relax.

III. GENERALIZATION OF THE ERGODIC THEORY PROOF OF RELAXATION TO AUTONOMOUS NON-HAMILTONIAN SYSTEMS

We now generalize this derivation so that it applies to any dynamics that preserves a mixing, equilibrium distribution $f_{eq}(\boldsymbol{\Gamma})$ for a finite system, as first presented in [2] (see Appendix A for definition of an equilibrium distribution). The phase space vector could be augmented by additional variables such as the Nosé-Hoover thermostat multiplier or the system volume to cover a variety of different systems (*e.g.* Nosé-Hoover thermostatted dynamics or Nosé-Hoover isothermal isobaric dynamics). However, we do not show this explicitly in our notation.

Using our definition of an equilibrium system we write the equilibrium distribution as,

$$
f_{eq,h}(\boldsymbol{\Gamma}) = \frac{\exp[-h(\boldsymbol{\Gamma})]}{\int_D d\boldsymbol{\Gamma} \, \exp[-h(\boldsymbol{\Gamma})]},
\tag{8}
$$

where $h(\boldsymbol{\Gamma})$ is some real, integrable function of the (possibly augmented) phase space vector, $\boldsymbol{\Gamma}$ defined over some phase space domain, D. We note that choice of this form of

---

[6] To extend to nonequilibrium steady states (NESS), we would need to carry out the calculation in terms of singular measures and ensure that the initial and final measures are absolutely continuous with respect to each other, as considered by Sinai [3].

[7] Note that $f(\boldsymbol{\Gamma}; 0)$ can be treated as a phase function because it has no explicit time-dependence, in contrast to $f(S^t\boldsymbol{\Gamma}; t)$ or $f(\boldsymbol{\Gamma}; t)$.

the potential is quite general and ensures that $f_{eq,h}(\mathbf{\Gamma}) > 0$ at any point in D. *All* positive smooth equilibrium distributions must be expressible in a form given by (8) (*e.g.* the canonical distribution or the isothermal isobaric distribution. In these two cases the appropriate dynamics which preserve these distributions are well known [6].). Equilibrium distributions must be autonomous because by definition the dissipation must be zero (Appendix A) and therefore they must be time independent.

We compute the average of some physical phase space function $A(\mathbf{\Gamma})$ at some time t with an initial distribution $f_0(\mathbf{\Gamma}) \neq f_{eq,h}(\mathbf{\Gamma})$ and undergoing dynamics that would preserve (8):

$$\langle A(t) \rangle_0 = \int_D d\mathbf{\Gamma}\, A(S^t \mathbf{\Gamma}) f_0(\mathbf{\Gamma})$$

$$= \frac{\int_D d\mathbf{\Gamma}\, A(S^t \mathbf{\Gamma}) f_0(\mathbf{\Gamma}) \exp[h(\mathbf{\Gamma})] \exp[-h(\mathbf{\Gamma})]}{\int_D d\mathbf{\Gamma}\, \exp[-h(\mathbf{\Gamma})]}$$

$$\times \int_D \exp[-h(\mathbf{\Gamma})]$$

$$= \langle A(S^t \mathbf{\Gamma}) f_0(\mathbf{\Gamma}) \exp[h(\mathbf{\Gamma})] \rangle_{eq,h} \int_D \exp[-h(\mathbf{\Gamma})]$$

$$\xrightarrow[t \to \infty]{} \langle A(\mathbf{\Gamma}) \rangle_{eq,h} \langle f_0(\mathbf{\Gamma}) \exp[h(\mathbf{\Gamma})] \rangle_{eq,h}$$

$$\times \int_D \exp[-h(\mathbf{\Gamma})]$$

$$= \langle A(\mathbf{\Gamma}) \rangle_{eq,h} \frac{\int_D d\mathbf{\Gamma}\, f_0(\mathbf{\Gamma}) \exp[h(\mathbf{\Gamma})] \exp[-h(\mathbf{\Gamma})]}{\int_D d\mathbf{\Gamma}\, \exp[-h(\mathbf{\Gamma})]}$$

$$\times \int_D d\mathbf{\Gamma}\, \exp[-h(\mathbf{\Gamma})]$$

$$= \langle A(\mathbf{\Gamma}) \rangle_{eq,h} \int_D d\mathbf{\Gamma}\, f_0(\mathbf{\Gamma}) = \langle A(\mathbf{\Gamma}) \rangle_{eq,h}, \tag{9}$$

where $\langle ... \rangle_{eq,h}$ denotes an average with respect to the mixing equilibrium distribution given by (8). We note that the partition function associated with $f_0(\mathbf{\Gamma})$ and $\int_D d\mathbf{\Gamma}\, \exp[-h(\mathbf{\Gamma})]$ must be finite and non-zero otherwise the derivation cannot be completed. We also require a form of ergodic consistency for all $\mathbf{\Gamma}$ such that if $f_0(\mathbf{\Gamma}) \neq 0$, $f_{eq,h}(\mathbf{\Gamma}) \neq 0$ which is guaranteed in (8).

These latter points mean that the derivation cannot be directly extended to nonequilibrium steady states because the relevant density functions do not exist. In this respect, we observe that Sinai's proof [3] can be extended to singular mixing distributions, but unlike the cases we usually consider in physics, the starting state must also be represented by a singular distribution. In our notation this system could be treated if D was considered to be the domain of the mixing state for which the density in the ostensible phase space

was singular, and the argument was reformulated in terms of measures instead of densities. However, for a dissipative steady state this domain would be a very complicated object of zero volume in the ostensible phase space.

Examples where application of (9) is used to show that phase averages relax to their equilibrium values, for non-Hamiltonian dynamics are given in Appendix C.

IV. EXPRESSION OF PHASE AVERAGES IN TERMS OF LOWER DIMENSIONAL DISTRIBUTION FUNCTIONS

Formally, the calculation of averages in (4) requires either knowledge of the full, time dependent, phase space distribution, $f(\mathbf{\Gamma}; t)$ or the initial phase space distribution. However, phase averages of properties that are phase functions of only, say, single particle or pairs of particles can be expressed *exactly* as functionals of the single-particle distribution function, $f^{(1)}(\mathbf{r}_1, \mathbf{p}_1; t)$, and the two-particle distribution, $f^{(2)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2; t)$, which may be obtained by integrating the full phase space distribution over N-1 or N-2 particle coordinates and momenta, respectively [7]. Of course in real systems there may be higher order interactions including 3 or even 4-body interactions, and these will also need a knowledge of the corresponding higher order distributions ($f^{(3)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3; t)$ and $f^{(4)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4; t)$, respectively). However, in a phase space with $O(6N_A)$ dimensions, where N_A is Avogadro's Number ($O(10^{23})$), these few-body interactions involve *vastly* lower dimensionality than is present in the full phase space. An example of a case where knowledge of just the single particle and pair distributions is informative is in a system where the potential energy of the entire system, $\Phi(\mathbf{\Gamma})$ can be written as a sum of pairwise interactions $\Phi(\mathbf{\Gamma}) = \sum_{i>j}^N \phi(\mathbf{r}_{ij})$ where $\phi(\mathbf{r}_{ij})$ is the potential energy of particles i and j which are separated by a distance $\mathbf{r}_{ij}$. In this case the average energy or pressure or stress in the system at time t, can be determined *exactly* from the singlet and pair distribution functions (the singlet for the kinetic terms and the pair distribution for the potential energy terms).

There is a further important observation to make about (4). Although we only need low order time dependent distributions to compute the time dependent averages of low order physical properties, in order to compute the time dependent single-particle and pair distributions, $f^{(1)}(\mathbf{r}_1, \mathbf{p}_1; t)$, $f^{(2)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2; t)$ does require the use of the full $6N$-dimensional phase continuity equation. Exact closed, equations of motion for $f^{(1)}(\mathbf{r}_1, \mathbf{p}_1; t)$, $f^{(2)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{p}_1, \mathbf{p}_2; t)$ do not exist. Instead one has the BBKGY hierarchy [8] where the equation of motion for a distribution of a given order is coupled to the next higher order distribution function.

Phase functions of single particles and pairs of particles, with a distribution function , can in general be written as,

$$\langle A \rangle_t = \frac{1}{\int_D d\mathbf{\Gamma}} \bullet \int_D d\mathbf{\Gamma} \, A(S^t\mathbf{\Gamma})f(\mathbf{\Gamma};0) \int_D d\mathbf{\Gamma}$$

$$= \left\langle A(S^t\mathbf{\Gamma})f(\mathbf{\Gamma};0) \right\rangle_{\mu c} \int_D d\mathbf{\Gamma}$$

$$= \left\langle A^{(1)}[S^t(r_1,\mathbf{p}_1)]f(\mathbf{\Gamma};0) \right.$$

$$+ \left. A^{(2)}[S^t(r_1,r_2,\mathbf{p}_1,\mathbf{p}_2)]f(\mathbf{\Gamma};0) \right\rangle_{\mu c} \int_D d\mathbf{\Gamma}$$

$$= \left\langle A^{(1)}[S^t(r_1,\mathbf{p}_1)]f^{(1)}(r_1,\mathbf{p}_1;0) \right.$$

$$+ \left. A^{(2)}[S^t(r_1,r_2,\mathbf{p}_1,\mathbf{p}_2)]f^{(2)}(r_1,r_2,\mathbf{p}_1,\mathbf{p}_2;0) \right\rangle_{\mu c} \int_D d\mathbf{\Gamma}$$

$$\tag{10}$$

where $A^{(1)}$ and $A^{(2)}$ refer to the single particle and pair particle contributions to A. In the last line for the single-particle term we can integrate over N-1 particle coordinates and momenta and write an equivalent expression involving a smooth single particle distribution function $f^{(1)}(\mathbf{r}_1,\mathbf{p}_1;0) \equiv \int d\mathbf{r}_2...d\mathbf{r}_N d\mathbf{p}_2...d\mathbf{p}_N f(\mathbf{r}_1,...,\mathbf{p}_N;0)$ and we can obviously do an analogous set of integrations for the two-body term but it is easier not to perform these integrations.

A few more words need to be said about $\langle A(S^t\mathbf{\Gamma})f(\mathbf{\Gamma};0)\rangle_{\mu c}$. This function is an equilibrium microcanonical, cross-time correlation function. It results from the fact that for Hamiltonian dynamics, any time dependent nonequilibrium ensemble average, say $\langle A \rangle_t$ (*i.e.* the ensemble average computed with respect to the evolved distribution), equals a time dependent nonequilibrium average $\langle A(S^t\mathbf{\Gamma})\rangle_0$ computed with respect to the initial distribution $f(\mathbf{\Gamma};0)$ (see equation (4)). Using the expansion in terms of low order distributions, for a phase variable that is a function on single and pair particle distributions,

$$\lim_{t\to\infty} \langle A \rangle_t = \left\{ \left\langle A^{(1)}(\mathbf{r}_1,\mathbf{p}_1) \right\rangle_{\mu c} \langle f(\mathbf{\Gamma};0)\rangle_{\mu c} \right.$$

$$+ \left. \langle A^{(2)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{p}_1,\mathbf{p}_2)\rangle_{\mu c}\langle f(\mathbf{\Gamma};0)\rangle_{\mu c} \right\} \int_D d\mathbf{\Gamma}$$

$$= \left\{ \langle A^{(1)}(\mathbf{r}_1,\mathbf{p}_1)\rangle_{\mu c} + \langle A^{(2)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{p}_1,\mathbf{p}_2)\rangle_{\mu c} \right\}$$

$$\left[\frac{\int_D d\mathbf{\Gamma} f(\mathbf{\Gamma};0)}{\int_D d\mathbf{\Gamma}} \right] \int_D d\mathbf{\Gamma} = \langle A(\mathbf{\Gamma})\rangle_{\mu c}.$$

$$\tag{11}$$

Therefore, although we expect that in a relaxing system that starts from a smooth nonequilibrium distribution, the full N-particle distribution will relax towards a highly structured (fractal) phase space distribution, convergence is only necessary for the distribution functions up to the same order

as the phase functions, in order for them to converge. However when we integrate over N-1 or N-2 particles *etc.*, the resulting low order, single-particle or two-body distribution functions will always be smooth. We will show that these low order distribution functions, under our stated conditions, relax towards equilibrium.

It has been proven that low order, steady state distribution functions can be written as ensemble averages of phase variables (see Appendix D). Therefore the mixing condition can be applied to them, just like any other phase function. An explanation of why ensemble averages of phase functions can converge when the distribution does not, can be obtained by noting that, by definition, the smoothness or degree of variation of a phase variable in phase space, does not change in time: trivially, it is just fixed. This is not the case for the N-particle distribution function, which continues to evolve and develop more and more structure with time. If, at some point in the future, the length-scale of the variations in the distribution function will be much smaller than the smallest length-scale over which the phase function of interest changes significantly, then after that point in time the value of the phase function will no longer change even though the full distribution function continues to develop even more fine-scaled structure. The less smooth the phase function is in phase space, the longer it will take for the phase function averages to become time independent. The possibility of nonuniformity in the decay of correlations, which would lead to differences in relaxation times, is also well established [9].

We also note that the degree of structure in the $f^{(i)}(\mathbf{r}_1,\ldots,\mathbf{r}_i,\mathbf{p}_1,\ldots\mathbf{p}_i;0)$ distribution function will be less than in $f^{(i+1)}(\mathbf{r}_1,\ldots,\mathbf{r}_{i+1},\mathbf{p}_1,\ldots\mathbf{p}_{i+1};0)$ due to integration of one particle's degrees of freedom, and therefore it is expected that the low order distribution functions (and hence the low order phase functions) will converge most quickly. In contrast, variables that are not phase variables and which at any time vary on the same (or smaller) length-scale as the N-particle distribution function (say $\ln[f(\mathbf{\Gamma};t)]$ – as appears in the expression for the entropy, and which is explicitly time-dependent; or $\delta(\mathbf{\Gamma} - \mathbf{\Gamma}_0)$, which is not a function) will, like $f(\mathbf{\Gamma};t)$, never converge to an equilibrium value.

The analysis presented above shows that if the initial distribution is not the equilibrium distribution and the system is mixing, then phase averages and low order distribution functions relax to their equilibrium values. Another way of obtaining this result is to use the dissipation theorem [10], which shows that in ΩT-mixing systems *all* ensemble averages of phase functions become time independent at long times. Since the low order distribution functions are phase functions (see Appendix D), they *must* relax to time independent distributions as a system relaxes, either towards a nonequilibrium steady state or towards an equilibrium state. This does *not* imply that the full phase space den-

sity relaxes to the equilibrium density. The proof based on the mixing property also carries no information about fundamental physical properties such as relaxation times that are fundamental, *e.g.* for the existence of transport coefficients.

The proof of relaxation to equilibrium using the mixing property of the steady state complements our previous proofs of relaxation to equilibrium in thermostatted or indeed barostatted systems that are ΩT-mixing. The proof based on mixing is very concise. However it must be supplemented with our recent definition of an equilibrium system. It also gives no information about relaxation timescales (which are important in the computation of the spectra of frequency dependent transport coefficients) and it gives no information in itself about the relaxation of distribution functions.

We know from the ΩT-mixing equilibrium relaxation theorems that the necessary and sufficient condition for the relaxation of averages, is that the system must be ΩT-mixing and that an equilibrium state exists. [Of course in contradistinction, the T-mixing condition is a sufficient but not necessary condition.] The mere existence of a mixing equilibrium state generated by the dynamics implies that the transient states are ΩT-mixing and that all systems possessing mixing equilibrium states will in fact relax towards these states at long times – at least as can be inferred from the calculation of averages of physical phase functions from initial states characterized by non-singular distributions. This is a somewhat surprising result.

As was the case for Hamiltonian dynamics, (9) proves the relaxation of averages towards their equilibrium values. We cannot know from this proof whether the full phase space distribution relaxes towards its equilibrium form or not. This is only established for the low order distribution functions by realising that the averages of physical quantities are exact functionals of only the low order terms in the Green's expansion of the full distribution function.

Acknowledgements

Appendix A.

A precise definition of equilibrium was given in our papers [4,5]. For instance if $f_{eq}(\mathbf{\Gamma})$ is a solution of the possibly thermostatted and/or barostatted Newtonian dynamics it is clear that, $\partial f_{eq}(\mathbf{\Gamma})/\partial t = 0$ but it is well known that in general $df_{eq}(\mathbf{\Gamma})/dt \neq 0$. If a distribution is even under the time reversal mapping, all fluxes vanish everywhere in the system

but this does not mean the system is an equilibrium distribution. The question of deciding when a system is in equilibrium or not, or indeed whether a specific dynamics is capable of preserving an equilibrium distribution, is in general, not completely straightforward.

From our equilibrium relaxation theorems we define an equilibrium system to be a combination of a smooth distribution function $f(\mathbf{\Gamma};t)$ and a dynamics $\dot{\mathbf{\Gamma}}(\mathbf{\Gamma}(t))$ for which the dissipation function is identically zero (almost) everywhere in the accessible phase space domain: $\Omega(\mathbf{\Gamma}) = 0, \quad \mathbf{\Gamma} \in D$. The instantaneous dissipation function is defined as [4]:

$$\Omega(S^t\mathbf{\Gamma}) \equiv$$
$$\equiv -\frac{\partial}{\partial \mathbf{\Gamma}} \bullet \dot{\mathbf{\Gamma}}(S^t\dot{\mathbf{\Gamma}}) - \dot{\mathbf{\Gamma}}(S^t\dot{\mathbf{\Gamma}}) \bullet \frac{\partial \ln(f(S^t\mathbf{\Gamma};0))}{\partial \mathbf{\Gamma}}. \quad (A.1)$$

For a specified initial phase space distribution the instantaneous dissipation function is just another (very important) phase function. The dissipation theorem confirms that if a system is sampled from an equilibrium distribution of states, that distribution must be preserved by the dynamics for all time [10,11].

It may be non-trivial to establish whether a system is in equilibrium or not. A clear example of this may be obtained by considering systems thermostatted by the so-called Gaussian μ-thermostats. Consider a system of interacting particles where the wall particles, which act as a thermal reservoir, are subject to a Gaussian μ-thermostat [12] where the equation of change for the βth Cartesian component particle momentum is, $\dot{p}_{\beta,i} = F_{\beta,i} - \alpha p^\mu_{\beta,i}\text{sgn}(p_{\beta,i})$, $F_{\beta,i}$ is the βth Cartesian component of the standard interatomic force felt by particle *i*, due to all the other particles in the system and a is chosen to fix say, the sum of the $(\mu+1)^{th}$ moment of the particle momenta. It turns out that equilibrium is only possible if $\mu = 1$. This is clearly understood using our definition of equilibrium. If $\mu \neq 1$ the system is autodissipative and equilibrium is not possible [12].

In our proof [4] of relaxation to equilibrium using the ΩT-mixing[8] property, we establish that a given density is an equilibrium density by showing it has no dissipation at any point in phase space. We show that the phase space continuity equation then implies that density will be preserved by the specified dynamics and the Dissipation Theorem implies that if the system is sampled according to this density, the phase variables will remain unchanged.

The ΩT-mixing condition requires that in the late time limit, say after time t_c, the values of phase averages no longer change, which means that the system has relaxed. We can then use the Dissipation Theorem to show that a perturbed system evolving according to equilibrium equations of motion (dynamics is capable of preserving an equilibrium distribution) relaxes to equilibrium. To do this, we assume that the phase averages are given by an ensemble average over a smooth distribution function for $t \geq$

---

[8] A ΩT-mixing system is one for which, $\lim_{t\to\infty} \int_0^t \langle \Omega(\mathbf{\Gamma})B(S^t\mathbf{\Gamma})\rangle_0 \, ds = L_A \in R$ for all phase variables, B and where L_A is a finite real number.

t_c, (say, $f'(\mathbf{\Gamma}) = e^{-\beta H(\mathbf{\Gamma})+g'(\mathbf{\Gamma})}/Z'$) and define a dissipation function associated with this distribution function and the dynamics of the system, Ω'. Using the dissipation theorem and the ΩT-mixing condition then, $\frac{d}{dt}\langle\Omega'(t)\rangle = \frac{d}{dt}\int_0^t \langle\Omega'(s)\Omega'(0)\rangle\,ds = 0$ for all $t \geq 0$, where the averages are with respect to the proposed distribution function of the evolved system. Evaluating this at $t = 0$ gives $\langle\Omega'(0)\Omega'(0)\rangle = 0$, and therefore the dissipation function must be zero at all points in phase space at this time if the system has relaxed to equilibrium. If the equations of motion are capable of preserving an equilibrium distribution then there will be a phase variable, g' that allows this condition to be satisfied.

Appendix B. Comparison of the infinitely thin shell and finite shell microcanonical distributions

Although the infinitely thin shell microcanonical distribution ($\delta E \to 0$) may be mixing and weak T-mixing[9], the finite δE system is not because every set of ensemble members with a specified energy never mixes with a set with a different energy. There are infinitely many subsets of phase space with distributions that have $\partial f(\mathbf{\Gamma})/\partial t = df(\mathbf{\Gamma})/dt = 0$. In the $\delta E \to 0$ limit, the density is equivalent to the energy hypersurface areal density, $\sigma(\mathbf{\Gamma})$ [13]:

$$\sigma(\mathbf{\Gamma}) = \frac{\left\|\dot{\mathbf{\Gamma}}\right\|^{-1}}{\int_{H(\mathbf{\Gamma})=E} d\mathbf{\Gamma}_\perp \left\|\dot{\mathbf{\Gamma}}\right\|^{-1}}, \quad \forall\,\mathbf{\Gamma}:H(\mathbf{\Gamma})=E,$$
(B.1)

($d\mathbf{\Gamma}_\perp$ denotes an infinitesimal element *in* the energy hypersurface). This expression for the density is obtained by realising that for the energy shell $\delta E = \delta\Gamma_\parallel \|\nabla H\| = \delta\Gamma_\parallel \left\|-\mathbf{J}\bullet\dot{\mathbf{\Gamma}}\right\| = \delta\Gamma_\parallel \left\|\dot{\mathbf{\Gamma}}\right\|$ where $\mathbf{J}$ is the usual symplectic matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\delta\Gamma_\parallel$ is the limiting shell thickness measured using some phase space distance metric along the gradient of the Hamiltonian, $\nabla H(\mathbf{\Gamma})$. $\delta\Gamma_\parallel$ is scalar and a function of $\mathbf{\Gamma}$, and $\lim_{\delta E\to 0}\int_{\mathbf{\Gamma}\in D(\delta E)} d\mathbf{\Gamma} = \int_{H(\mathbf{\Gamma})=E} d\mathbf{\Gamma}_\perp \delta\Gamma_\parallel(\mathbf{\Gamma})$.

Appendix C. Examples of how mixing leads to relaxation of phase averages to their equilibrium values

As examples of the application of equation (9), we consider two possibilities here. In the first case we consider an initial distribution of states that is consistent with a canonical equilibrium system at temperature T_0 where $\beta_0 = 1/k_B T_0$ (here k_B is Boltzmann's constant), $f_0(\mathbf{\Gamma}) = e^{-\beta_0 H(\mathbf{\Gamma})}/\int_D d\mathbf{\Gamma}\,e^{-\beta_0 H(\mathbf{\Gamma})}$. At time zero, a Nosé-Hoover dynamics with a thermostat of temperature T_1, and $\beta_1 = 1/k_B T_1$ is applied. Nosé-Hoover thermostatted dynamics of many-particle systems described by realistic interaction potentials is mixing. That is,

$$\lim_{t\to\infty} \langle B(t)A(0)\rangle_1 = \langle B(t)\rangle_1 \langle A(0)\rangle_1 \qquad \text{(C.1)}$$

where $\langle\ldots\rangle_1$ is the ensemble average with respect to the canonical distribution with a temperature T_1, and the dynamics is Nosé-Hoover thermostatted dynamics with a thermostat of temperature T_1. This dynamics preserves that canonical distribution, in which case at all times,

$$\langle B(t)\rangle_1 = \langle B(0)\rangle_1. \qquad \text{(C.2)}$$

Let us consider a typical property, say the microscopic contribution to the pressure $p(\mathbf{\Gamma})$. Initially the average pressure of the system will be that of the system at temperature T_0, but if the points in phase space are propagated with Nosé-Hoover dynamics and a thermostat of temperature T_1, we can follow (9) to show that at some time later:

$$
\begin{aligned}
\langle p(t)\rangle_0 &= \frac{\int_D d\mathbf{\Gamma}\,p(S^t\mathbf{\Gamma})e^{-\beta_0 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma}\,e^{-\beta_0 H(\mathbf{\Gamma})}}\\
&= \frac{\int_D \mathbf{\Gamma}p(S^t\mathbf{\Gamma})e^{-\beta_0 H(\mathbf{\Gamma})+\beta_1 H(\mathbf{\Gamma})-\beta_{1??}H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma}e^{-\beta_1 H(\mathbf{\Gamma})}}\\
&\quad \int_D d\mathbf{\Gamma}e^{-\beta_1 H(\mathbf{\Gamma})}\\
&= \langle p(t)e^{-\beta_0 H(0)+\beta_1 H(0)}\rangle_1 \frac{\int_D d\mathbf{\Gamma}e^{-\beta_1 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma}e^{-\beta_0 H(\mathbf{\Gamma})}}
\end{aligned}
$$
(C.3)

In this case, the domain, *D*, is all values of the momenta of each particle and the all positions, given periodic boundary conditions $[\mathbf{0}, \mathbf{L}]$. The dynamics used to evolve the pressure is the Nosé-Hoover dynamics with a thermostat of temperature T_1. Since this dynamics is mixing, then using (C.1), $\lim_{t\to\infty}\langle p(t)e^{-\beta_0 H(0)+\beta_1 H(0)}\rangle_1 = \langle p(t)\rangle_1 \langle e^{-\beta_0 H(0)+\beta_1 H(0)}\rangle_1$ and using (C.2), $\langle p(t)\rangle_1 = \langle p(0)\rangle_1$. Therefore, taking the long time limit (C.3) becomes:

---

[9] A weakly T-mixing system is one where $\lim_{t\to\infty}\left[\langle A(\mathbf{\Gamma})B(S^t\mathbf{\Gamma})\rangle_0 - \langle A(\mathbf{\Gamma})\rangle_0 \langle B(S^t\mathbf{\Gamma})\rangle_0\right] = 0$. Here A and B are phase variables, and the ensemble averages are over the initial measure. This condition differs from the mixing condition, which refers to the invariant measure rather than the initial measure. However in the case that the system starts in the equilibrium distribution that is preserved by the dynamics, they become equivalent. The strong form of T-mixing ensures that the correlations not only decay, but decay quickly enough so that $\lim_{t\to\infty}\int_0^t \langle A(\mathbf{\Gamma})B(S^t\mathbf{\Gamma})\rangle_0 - \langle A(\mathbf{\Gamma})\rangle_0 \langle B(S^t\mathbf{\Gamma})\rangle_0\,ds = L_A \in R$. The weak form of T-mixing is sufficient for the cases discussed in this paper.

$$\lim_{t\to\infty} \langle p(t) \rangle_0 = \langle p(0) \rangle_1 \left\langle e^{-\beta_0 H(\mathbf{0}) + \beta_1 H(0)} \right\rangle_1$$

$$\frac{\int_D d\mathbf{\Gamma}\, e^{-\beta_1 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma}\, e^{-\beta_0 H(\mathbf{\Gamma})}}$$

$$= \langle p(0) \rangle_1 \frac{\int_D d\mathbf{\Gamma} e^{-\beta_0 H(\mathbf{0}) + \beta_1 H(0)} e^{-\beta_1 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}} \qquad \text{(C.4)}$$

$$\frac{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma} e^{-\beta_0 H(\mathbf{\Gamma})}}$$

$$= \langle p(0) \rangle_1 \frac{\int_D d\mathbf{\Gamma} e^{-\beta_0 H(\mathbf{0})}}{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}} \frac{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma} e^{-\beta_0 H(\mathbf{\Gamma})}}$$

$$= \langle p(0) \rangle_1$$

This shows that in the long time limit, the pressure will become the equilibrium pressure of the canonical distribution with at temperature T_1. The only condition required was that the dynamics was mixing.

Another example would be if in the initial state, the distribution was a subset of the canonical ensemble, with all x-positions of the particles fixed to be within the range $[0, \ L_x/2]$. We can write that distribution as,

$$f(\mathbf{\Gamma}) = \frac{e^{-\beta_1 H(\mathbf{\Gamma})} S(q_x)}{\int_D d\mathbf{\Gamma}\ e^{-\beta_1 H(\mathbf{\Gamma})} S(q_x)} \qquad \text{(C.5)}$$

where $S(q_x) = \prod_{i=1} (\Theta(q_{xi}) + \Theta(-q_{xi} + L_x/2) + 1)$ is a switch that is 1 when all q_{xi} are in the range $[0, \ L_x/2]$, and zero otherwise and Θ is the Heaviside-step function. Now consider the dynamics be Nosé-Hoover thermostatted dynamics with a thermostat of temperature T_1. Then equations (C.1) and (C.2) will apply and it can be shown, that for a property such as the pressure, $p(\mathbf{\Gamma})$:

$$\lim_{t\to\infty} \langle p(t) \rangle_0 = \lim_{t\to\infty} \frac{\int_D d\mathbf{\Gamma}\, p(S^t\mathbf{\Gamma}) e^{-\beta_1 H(\mathbf{\Gamma})} S(q_x)}{\int_D d\mathbf{\Gamma}\ e^{-\beta_1 H(\mathbf{\Gamma})} S(q_x)}$$

$$= \lim_{t\to\infty} \frac{\int_D d\mathbf{\Gamma} p(S^t\mathbf{\Gamma}) S(q_x) e^{-\beta_1 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})} S(q_x) \int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}}$$

$$\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}$$

$$= \lim_{t\to\infty} \langle p(t) S(q_x(0)) \rangle_1 \frac{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})} S(q_x)}$$

$$= \langle p(0) \rangle_1 \langle S(q_x(0)) \rangle_1 \frac{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})} S(q_x)}$$

$$= \langle p(0) \rangle_1 \frac{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})} S(q_x)}{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}} \frac{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})}}{\int_D d\mathbf{\Gamma} e^{-\beta_1 H(\mathbf{\Gamma})} S(q_x)}$$

$$= \langle p(0) \rangle_1. \qquad \text{(C.6)}$$

In this case, although the initial distribution is a step function and this distribution is not preserved by any dynamics in D, it is not the properties of the initial distribution that

are important in the results, but the properties of the equilibrium distribution that the system is relaxing to.

Appendix D. Computation of low order distribution functions as averages of phase functions

In a nonequilibrium system with pair interactions only, the configurational component of the pressure tensor and the internal energy are functionals of the anisotropic pair distribution function $g(\mathbf{r})$, only. All higher order distributions have *no* effect on these quantities. The pair distribution is defined as: given one particle is at the origin, what is the probability per unit volume that a second particle is located at a position r away from the origin.

If the system is in a time stationary state, the pair distribution may be expanded as a time independent spherical harmonic expansion:

$$g(\mathbf{r}) = \left\langle \sum_{m,n} A_{m,n}(r) Y_{m,n}^e(\theta, \phi) + B_{m,n}(r) Y_{m,n}^o(\theta, \phi) \right\rangle, \qquad \text{(D.1)}$$

where

$$Y_{m,n}^e = \cos(m\phi) P_n^m(\cos\theta),$$
$$Y_{m,n}^o = \sin(m\phi) P_n^m(\cos\theta) \qquad \text{(D.2)}$$

are the usual polar and azimuthal angles. Here we follow the notation of Vol II. of Morse and Feshbach [14].

The expansion coefficients can be obtained from ensemble averages of appropriate phase functions:

$$A_{m,n}(r) = \frac{\varepsilon_m (2n+1)(n-m)!}{4\pi\rho(n+m)! r^2} \left\langle Y_{m,n}^e(\theta, \phi) \right\rangle_{r, r+dr}$$

$$B_{m,n}(r) = \frac{\varepsilon_m (2n+1)(n-m)!}{4\pi\rho(n+m)! r^2} \left\langle Y_{m,n}^o(\theta, \phi) \right\rangle_{r, r+dr}$$

$$\text{(D.3)}$$

where $\varepsilon_m = 1 : m = 0; \ \varepsilon_m = 2 : m \neq 0$ and $\langle ... \rangle_{r, r+dr}$ denotes an ensemble average taken over a thin shell of radius r. In (D.3) $3N$ momenta have been integrated out as well as $3(N-1)$ coordinates leaving only the relative positions of pair of particles expressed in polar coordinates. These averages are computed in the nonequilibrium steady state where the pair distribution function is time independent.

Averages of physical properties can then be obtained by integrating the appropriate expansion coefficients over the pair separation. For instance in a fluid undergoing steady planar Couette shear flow, $\dot{\gamma} = \partial u_x/\partial y$, the configurational component of the xy-element of the ensemble averaged pressure tensor, $\langle P_{xy}^\phi \rangle$ is given exactly by

$$\langle P_{xy}^\phi \rangle = \frac{-2\pi\rho^2}{5} \int_0^\infty dr\ B_{22}(r) r^3 \frac{du(r)}{dr}, \qquad \text{(D.4)}$$

where $u(r)$ is the potential energy of two atoms a distance r apart (pair interactions only are assumed). This shows

how low order distribution functions like the pair distribution $g(r)$ can be expressed as ensemble averages of phase functions, as in (D.3).

We note that even in a system with pair interactions only, higher order thermodynamic quantities like specific heats will be functionals of the 1, 2, and 3-body distribution functions. So in systems with a given order for potential energy interactions the order of the distribution function that is needed to calculate thermodynamic averages, is property dependent.

In nonequilibrium steady states the N-particle distribution function approaches a fractal in the long time limit [15]. In macroscopic systems that obey Navier-Stokes hydrodynamics and which are non turbulent, the Kaplan-Yorke dimension [16] of the steady state attractor is only ~10 less than the ostensible dimension of phase space (6 times Avogadro's Number – ~6×10^{23}) [15]. All distributions that are of lower order than the Kaplan-Yorke dimension are expected to be smooth. This means that distributions in such systems are smooth up to a very high dimension indeed. However any dimensional reduction no matter how small of course makes the entropy ill defined since the entropy, unlike other thermophysical properties, is a functional of the full N-particle phase space distribution.

References

[1] D.J. Evans, S.R. Williams, L. Rondoni, *A mathematical proof of the zeroth "law" of thermodynamics and the nonlinear Fourier "law" for heat flow*, J. Chem. Phys., **137**, 194109 (2012).

[2] D.J. Evans, D.J. Searles, S.R. Williams, *Fundamentals of classical statistical thermodynamics, dissipation, relaxation and fluctuation theorems*, Wiley-VCH Verlag, Weinheim, Germany 2016.

[3] Y.G. Sinai, *Introduction to ergodic* theory, Princeton University Press, Princeton, NJ, USA 1976.

[4] D.J. Evans, D.J. Searles, S.R. Williams, *Dissipation and the relaxation to equilibrium*, J. Stat. Mech. Theor. Exp. **P07029** (2009).

[5] D.J. Evans, D.J. Searles, S.R. Williams, *A simple mathematical proof of Boltzmann's equal a priori probability hypothesis* in: C. Chmelik, N. Kanellopoulos, J. Karger, T. Doros (ed.) *Diffusion fundamentals III*, Leipziger Universitatsverlag, Leipzig, p. 367- 374, 2009.

[6] D.J. Evans G.P. Morriss, *Statistical mechanics of non-equilibrium liquids*, Academic, London 1990.

[7] H. Green, *The molecular theory of fluids*, North-Holland, Amsterdam 1952.

[8] R. Balescu, *Equilibrium and non equilibrium statistical mechanics*, Wiley Interscience, New York 1975.

[9] J.D. Crawford J.R. Cary, *Decay of correlations in a chaotic measure-preserving transformation*, Physica D **6**, 223 (1983).

[10] D.J. Evans, D.J. Searles, S.R. Williams, *On the fluctuation theorem for the dissipation function and its connection with response theory*, J. Chem. Phys. **128**, 014504 (2008).

[11] D.J. Evans, D.J. Searles, S.R. Williams, Erratum: *On the fluctuation theorem for the dissipation function and its connection with response theory* (vol 128, artn 014504, 2008), J. Chem. Phys. **128**, 249901 (2008).

[12] D.J. Evans, D.J. Searles, S.R. Williams, *Musings on thermostats*, J. Chem. Phys. **133**, 104106 (2010).

[13] C. Thompson, *Mathematical statistical mechanics*, Collier Macmillan Ltd, London 1972.

[14] P.M. Morse and H. Feshback, *Methods of theoretical physics*, McGraw-Hill, New York 1953.

[15] D.J. Evans, E.D. G. Cohen, and G.P. Morriss, *Viscosity of a simple fluid from its maximal lyapunov exponents*, Phys. Rev. A **42**, 5990 (1990).

[16] J.L. Kaplan and J.A. Yorke, *Chaotic behavior of multidimensional difference equations* in: H.O. Peitgen, and H.O. Walther (ed.) *Functional differential equations and approximation of fixed points*, Springer, Heidelberg, p. 204-227, 1979.

Denis J. Evans is an emeritus professor in the Department of Applied Mathematics, Australian National University (ANU), Canberra, Australia. He is also an Honorary Professor at The University of Queensland, Brisbane, Australia. Having obtained his PhD from ANU, he worked at Oxford and Cornell before returning to ANU. Later he held a Fulbright Fellowship at NIST, Boulder Colorado before taking a tenured position in the Research School of Chemistry ANU. Professor Evans has published over 300 scientific publications and two research monographs. He has been awarded numerous prizes including the Boys-Rahman Lectureship of the Royal Society of Chemistry (London) and the David Craig Medal of the Australian Academy of Science. He has been a fellow of that Academy since 1994. He was recently appointed a Member of the Order of Australia.

Stephen R. Williams completed his PhD in Physics at RMIT University, then went on to be a postdoctoral fellow at Utrecht University and a fellow at the Australian National University, Australia.

Lamberto Rondoni is professor of mathematical physics in the Department of Mathematical Sciences, Politecnico di Torino, Italy. Currently he coordinates the joint PhD program in pure and applied mathematics of Politecnico di Torino and Università di Torino. He obtained his PhD from Virginia Polytechnic Institute and State University in 1991. He held several postdoctoral positions in the USA before moving to Australia for a research associate position at the University of New South Wales. In 1995 he returned to Italy, where he became full professor in 2011.

Debra J. Searles (Bernhardt) is Senior Group Leader and Professor at the Australian Institute for Bioengineering and Nanotechnology (AIBN) and School of Chemistry and Molecular Biosciences at The University of Queensland. She is Director of the AIBN Centre for Theoretical and Computational Molecular Science. Debra obtained her PhD at The University of Newcastle, Australia, and held research positions at The University of Basel and the Australian National University. Before joining The University of Queensland in 2012, she was at Griffith University for 12 years.