ON TYPICALITY IN NONEQUILIBRIUM STEADY STATES

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

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

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

[3]Australian Institute for Bioengineering & Nanotechnology Centre for Theoretical and Computational Molecular Science, The University of Queensland, St Lucia, QLD 4072, Australia

[4]School of Chemistry and Molecular Biosciences, The University of Queensland, St Lucia, QLD 4072, Australia

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

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

[7]MICEMS, Universiti Putra Malaysia, 43400 Serdang Selangor, Malaysia

ABSTRACT

From the statistical mechanical viewpoint, relaxation of macroscopic systems and response theory rest on a notion of *typicality*, according to which the behavior of single macroscopic objects is given by appropriate ensembles: ensemble averages of observable quantities represent the measurements performed on single objects, because *"almost all"* objects share the same fate. In the case of non-dissipative dynamics and relaxation toward equilibrium states, *"almost all"* is referred to invariant probability distributions that are absolutely continuous with respect to the Lebesgue measure. In other words, the collection of initial micro-states (single systems) that do not follow the ensemble is supposed to constitute a set of vanishing, phase space volume. This approach is problematic in the case of dissipative dynamics and relaxation to nonequilibrium steady states, because the relevant invariant distributions attribute probability 1 to sets of zero volume, while evolution commonly begins in equilibrium states, *i.e.* in sets of full phase space volume.

We consider the relaxation of classical, thermostatted particle systems to nonequilibrium steady states. We show that the dynamical condition known as ΩT-mixing is *necessary* and *sufficient* for relaxation *of ensemble averages* to steady state values. Moreover, we find that the condition known as weak T-mixing applied to smooth observables is *sufficient* for ensemble relaxation to be independent of the initial ensemble. Lastly, we show that weak T-mixing provides a notion of typicality for dissipative dynamics that is based on the (non-invariant) Lebesgue measure, and that we call *physical ergodicity*.

I. INTRODUCTION

The study of the relaxation of systems made of many microscopic constituents obeying deterministic time-reversible equations of motion to stationary states, is long and celebrated; it finds its modern roots in Ludwig Boltzmann's Kinetic Theory. Boltzmann's approach was based on a notion that he called ergodicity [1, 2], which in the early XXth century was considered by physicists to be verified for the Hamiltonian systems of interest [3].

The mathematical notion of ergodicity that is commonly used today is that of metric transitivity, and in Hamiltonian dynamics (which preserve phase space volumes) it requires that almost all trajectories densely explore the phase space, $\mathcal{M}$ say; the trajectories that do not densely explore $\mathcal{M}$ constitute a set of vanishing phase space volume. This implies that (infinite) time averages of observables are equal to appropriate equilibrium phase space averages, apart from a set of vanishing phase space volume (or zero Lebesgue measure). Roughly, this means that *"typically"* the ensemble members enjoy the same behavior, *i.e.* those that do not constitute a set of vanishing Lebesgue measure.[1]

While the notion of ergodicity based on invariant measures led to many insights of physical interest [4-10], the idea of typicality that it entails is problematic in the case of dissipative dynamics, such as those of nonequilibrium molecular dynamics (NEMD) models. Indeed the invariant measures of such models are singular with respect to the Lebesgue measure, as they attribute probability 1 to sets of

[1] We refer to the time averages as to "observable values", because measurements of macroscopic quantities take a time that is large compared to the microscopic characteristic times, and within their accuracy they yield the result of very many (ideally infinitely many) interactions between measurement tool and system of interest. In phase space, this is commonly represented by an (ideally infinite) time average along a given phase space trajectory. Averages over insufficiently long times may yield sensibly different values, depending on the systems at hand. In particular, short time averages concerning systems of a small number of particles tend to be widely dispersed.

vanishing volume. Therefore, in the sense of these invariant measures, *"typically"* means that the exceptions may occupy a set whose volume equals the volume of the entire phase space, or that the set of ensemble members that are guaranteed to behave in the same way are almost none, from the point of view of the phase space volume. Considering that standard experiments leading to nonequilibrium steady states (NESS) start in equilibrium states, *i.e.* with initial conditions in the entire phase space, this notion of typicality appears too limited for the statistical mechanical interpretation of relaxation of dissipative systems.

Our main results are the following: for dissipative dynamics we determine conditions for relaxation of averages of physical properties, to steady state values for both *ensembles* and single experiments.

In particular, we show that the condition known as ΩT-mixing is *necessary* and *sufficient* to prove relaxation of the ensemble averages of physical properties to steady state values, while usually only sufficient conditions are given. We find that the condition known as weak T-mixing (wT-mixing) applied to smooth observables is sufficient for relaxation to be independent of the initial ensemble. In turn, wT-mixing for integrable functions makes relaxation independent of the initial phase point, apart from a negligible (zero volume) set of points, in accord with experimental observations in thermodynamic systems.

This result bridges for dissipative systems the gap between standard results in response theory and observations, in the sense that typicality is referred to the (non-invariant) Lebesgue measure, *i.e.* to the whole phase space and not just to a set of invariant probability 1. This is what we call *physical ergodicity*.

II. ERGODICITY AND MIXING

A stationary state is defined as one for which ensemble averaged physical properties does not change with time to within the measureable accuracy. This stationary state is typically formed by evolution from some initial (equilibrium) state for a transient period until the properties are no longer observed to change. If the stationary state is out of equilibrium, we have a NESS. For such a system, let the initial state be described by a probability distribution function f_0 on $\mathcal{M}$, and let the evolution of the physical microstates be represented by equations of motion on $\mathcal{M}$. Then, in contrast with the behaviour of averages of physical properties, unless the system is at equilibrium, the distribution function evolves from the initial transient f_0, taking a different form f_t at every time *t*, and never stopping its evolution. For example the Gibbs entropy of a NESS diverges at a constant rate towards negative infinity; ensemble averages of phase variables calculated with f_t may nevertheless converge.

In dynamical systems theory a steady state is associated with a construct that is represented by an invariant probability distribution (an invariant measure, in mathematical terms)[2] that does not need be associated with a probability density function when examined in ostensible phase space. If the dynamics are conservative, and the system is at equilibrium, the ensemble amounts to a probability density function f_0, but for dissipative systems, the ensemble has a singular distribution of phase points in ostensible phase space. In the equilibrium case, the dynamical systems terminology refers to invariant measures that are absolutely continuous with respect to the Lebesgue measure; in the case of dissipative systems it refers to singular invariant

[2] An invariant probability distribution would give phase space averages of all properties that are time-invariant to all limits of accuracy. If this measure has a density (an equilibrium system), then the density itself does not change with time. A probability distribution in phase space may be viewed as a collection of non-interacting objects in a given state, called an "ensemble".

measures. This means that, in this dynamical systems framework, a NESS corresponds to a lower dimensional subset $\mathscr{A}$ of $\mathcal{M}$, which has steady state probability 1 but zero probability with respect to f_0, because its phase space volume vanishes. Therefore, initial points in $\mathcal{M}$ will not lie on $\mathscr{A}$, apart from those in this invariant set of zero volume, and their time evolution will result in the never ending collapse of the phase space probability distribution *towards* a singular measure that attributes probability 1 to $\mathscr{A}$. Following Milnor [11], we refer to the invariant set, $\mathscr{A}$, as the NESS attractor. For a system that reaches a steady state, at sufficiently long times t the averages calculated with the invariant measure and with the evolved density f_t are equal to within some accuracy.[3]

Let us recall a few facts about ergodic theory. Consider a deterministic evolution $S^t : \mathcal{M} \to \mathcal{M}$, with notation meaning that $S^t\mathbf{\Gamma} \in \mathcal{M}$ represents the phase at time t along a trajectory starting at $\mathbf{\Gamma} \in \mathcal{M}$. The first important feature of ergodicity is the use of phase space averages to express the infinite time average of an observable $O : \mathcal{M} \to \mathbf{R}$, mathematically represented by,

$$\bar{O}(\mathbf{\Gamma}) = \lim_{t\to\infty} \bar{O}(\mathbf{\Gamma}; t) \equiv \lim_{t\to\infty} \frac{1}{t} \int_0^t O\left(S^s\mathbf{\Gamma}\right) \mathrm{d}s \tag{1}$$

for a system whose initial microstate is $\mathbf{\Gamma} \in \mathcal{M}$. Note that in writing (1) we are assuming the long time limit exists. We are therefore assuming that at long times a state is reached in which the time averages of physical observables are time stationary. In an ergodic system, it is postulated that for almost all initial phases $\mathbf{\Gamma}$

$$\bar{O}(\mathbf{\Gamma}) = \int_{\mathcal{M}} O(\mathbf{\Gamma}) \mathrm{d}\mu(\mathbf{\Gamma}) = \langle O \rangle_\mu \tag{2}$$

[3] Of course, these zero volume attractors are purely geometric entities, whose connection with observable features of physical systems is not direct.

where, $\langle\cdot\rangle_\mu$ denotes the phase space average with respect to an appropriate probability distribution μ on $\mathcal{M}$. Because of the limit in time, the averages $\langle O\rangle_\mu$ are *invariant* under the dynamics S^t. Under metric transitivity, Eq. (2) holds for almost all (with respect to μ) $\mathbf{\Gamma}\in\mathcal{M}$.

There are two equivalent formulations of mixing which, like ergodicity, refer to an invariant measure μ, and which are not restricted to conservative dynamics. Given two (μ- measurable) sets E and F in $\mathcal{M}$, mixing is verified if:

$$\left[\mu\left(E\cap S^tF\right)-\mu(E)\mu(F)\right]=\left[\mu\left(S^{-t}E\cap F\right)-\mu(E)\mu(F)\right]\xrightarrow[t\to\infty]{}0 \qquad (3)$$

where for a generic measurable set $D\subset\mathcal{M}$, $\mu(D)=\int_D d\mu$ is the probability of D, and $\mu(\mathcal{M})=\int_{\mathcal{M}} d\mu=1$ by definition. If μ has a density f, one may write $\mu(D)=\int_D f(\Gamma)d\Gamma$. Equation (3) means that the fraction of the set that was initially in F and is found in any other set E at large times, equals the probability of being in F. This happens if the evolving set S^tF spreads all over the region of $\mathcal{M}$ concerning the steady state. Equation (3) is equivalent to the following decay of correlations for all integrable observables A and B:

$$\left[\int_{\mathcal{M}} A(S^t\mathbf{\Gamma})B(\mathbf{\Gamma})\,\mathrm{d}\,\mu(\mathbf{\Gamma})-\int_{\mathcal{M}} A(\mathbf{\Gamma})\,\mathrm{d}\,\mu(\mathbf{\Gamma})\int_{\mathcal{M}} B(\mathbf{\Gamma})\,\mathrm{d}\,\mu(\mathbf{\Gamma})\right]$$

$$=\left[\left\langle\left(A\circ S^t\right)B\right\rangle_\mu-\langle A\rangle_\mu\langle B\rangle_\mu\right]\equiv C_\mu(t)\xrightarrow[t\to\infty]{}0 \qquad (4)$$

where $A\circ S^t$ represents the composition of the function A with the time evolution, *i.e.* $A\circ S^t\mathbf{\Gamma}\equiv A\left(S^t\mathbf{\Gamma}\right)\equiv A\left(\mathbf{\Gamma}(t)\right)$, and we have defined the correlation function C_μ.[4]

[4] The equivalence of condition (3) for all measurable sets and condition (4) for all integrable functions can be understood as follows. In the first place, both Eq. (3) and Eq. (4) imply a loss of memory about the initial conditions: Eq. (3) says that the points found in E at time zero could have come from

This result has been used to show that averages over an appropriate initial distribution approach those taken over the invariant mixing measure that is preserved by the dynamics [13]. For instance, let the initial phase space distribution of the system be given by a density $f_0(\mathbf{\Gamma})$, and let the distribution at time t be given by $f_t(\mathbf{\Gamma})$. Suppose further that the Lebesgue measure is mixing for the dynamics. One can then write:

$$\begin{aligned}\langle A\rangle_t &= \int_{\mathcal{M}} A(\mathbf{\Gamma}) f_t(\mathbf{\Gamma})\,\mathrm{d}\mathbf{\Gamma} \\ &= \int_{\mathcal{M}} A\left(S^t\mathbf{\Gamma}\right) f_0(\mathbf{\Gamma})\,\mathrm{d}\mathbf{\Gamma} \qquad (5) \\ &= \left\langle \left(A\circ S^t\right) f_0\right\rangle_{\mu_{mc}}\end{aligned}$$

where the second equality is due to the equivalence of the Heisenberg and Schrödinger representations of phase space averages [7], and $\langle\cdot\rangle_t$ denotes the average with respect to μ_t. Then using the mixing condition (4),

$$\lim_{t\to\infty}\left\langle \left(A\circ S^t\right) f_0\right\rangle_{\mu_{mc}} = \langle A\rangle_{\mu_{mc}}\langle f_0\rangle_{\mu_{mc}} = \langle A\rangle_{\mu_{mc}} \qquad (6)$$

where the final equality holds because $\langle f_0\rangle_{\mu_{mc}} = \int f_0(\mathbf{\Gamma})d\mathbf{\Gamma} = 1$, by definition of probability density. So for systems that are mixing with respect to μ_{mc} (which is preserved by the dynamics): $\lim_{t\to\infty}\langle A\rangle_t = \langle A\rangle_{\mu_{mc}}$. The conclusion is that ensemble

---

everywhere else in the phase space; Eq. (4) says that the correlation between any two observables is lost in time. Secondly, observe that $\mu(E)$ equals $\int_E \mathrm{d}\mu = \int_{\mathcal{M}} \chi_E(\Gamma)\mathrm{d}\mu(\Gamma)$, where the characteristic function χ is defined by $\chi_E(\Gamma)=1$ if $\Gamma\in E$ and $\chi_E(\Gamma)=0$ if $\Gamma\notin E$. It follows that Eq. (3) holds if correlations between integrable functions, which include χ_E, decay. At the same time, linear combinations of characteristic functions approximate to arbitrary precision integrable functions, therefore the validity of Eq. (3) can be used to imply Eq. (4). [12]

---

averages of observables converge from their initial values $\langle A \rangle_0$ to their asymptotic values $\langle A \rangle_{\mu_{mc}}$ corresponding to the uniform distribution in phase space [13, 14]. Because mixing implies ergodicity, if the ensemble averages in the long time limit converge, then time averages commencing from a point on $\mathcal{M}$ will also converge, again apparently implying relaxation.

This reasoning can be extended to other invariant densities [13, 15], as well as to dissipative dynamics and singular measures [13]. However, in the latter case it only refers to sets of zero phase space volume, because the evolving and the stationary measures have to be absolutely continuous with respect to each other.

III. THE DISSIPATION FUNCTION AND ΩT-MIXING

In the present paper we firstly use a notion recently introduced and known as ΩT-mixing, in order to study the relaxation of ensembles toward NESS [14, 16-18]. Like equilibrium states, NESS have stationary, time independent averages of phase functions (*e.g.* pressure, stress, energy *etc.*). However unlike equilibrium states, in dynamical systems theory the NESS of the dissipative dynamics is characterized by an invariant measure that attributes positive probability to sets of dimension lower than the ostensible dimension of phase space. [5] With dissipative dynamics, the phase space $\mathcal{M}$ may contain more than one NESS attractor (*i.e.* in Milnor's sense [11]), since they occupy only a vanishing volume.

To investigate relaxation both at and away from equilibrium, one can rely on a quantity introduced in 2000, namely the Dissipation Function Ω [19, 20]. Given a dynamical system $\dot{\mathbf{\Gamma}} = \mathbf{G}(\mathbf{\Gamma})$ and an initial distribution f_0 on $\mathcal{M}$, the dissipation function integrated over the time interval [*0 ,t*] is a phase variable defined by:

$$\Omega_{0,t}^{(f_0)}(\mathbf{\Gamma}) = \int_0^t \Omega^{(f_0)}\left(S^u\mathbf{\Gamma}\right)\mathrm{d}u = \ln\frac{f_0(\mathbf{\Gamma})}{f_0\left(\mathrm{M}^T S^t\mathbf{\Gamma}\right)} - \Lambda_{0,t}(\mathbf{\Gamma}) \tag{7}$$

where M^T is the time reversal map, *e.g.* $M^T\mathbf{\Gamma} \equiv (\mathbf{q}_1,...\mathbf{q}_N,-\mathbf{p}_1,...,-\mathbf{p}_N)$ for systems of *N* point particles, $\Lambda = \mathrm{div}\,\mathbf{G}$ is the phase space volume variation rate, and for every observable *A* the subscripts $0,t$ denote integration from time 0 to time *t* along the trajectory passing in $\mathbf{\Gamma}$ at time 0:

$$A_{0,t}(\mathbf{\Gamma}) = \int_0^t A\left(S^u\mathbf{\Gamma}\right)\mathrm{d}u\ . \tag{8}$$

[5] Because of this lower dimensionality, functions of the phase space probability density, such as the Gibbs entropy, are ill defined in a NESS. Alternatively, if followed in its evolution under dissipative dynamics from an initial equilibrium state, the Gibbs entropy diverges at a constant average rate [7]. This can also be seen as a consequence of the divergence to positive infinity of the probability density, observed from almost any point, combined with the shrinking volume of the occupied phase space.

Usually, the distribution f_0 is not invariant and changes in time, but Ω remains a phase variable, because Eq. (7) only refers to the distribution at a single time. As stressed by the notation in Eq. (7), Ω depends on f_0, however, for sake of simplicity, we will omit the superscript f_0 when there is no danger of confusion. Dividing $\Omega_{0,t}(\mathbf{\Gamma})$ by t and relying on the continuity in time of $f_0(S^t\mathbf{\Gamma})$ and of $\Lambda(S^t\mathbf{\Gamma})$, the instantaneous value Ω of the dissipation function is obtained from its integral representation (7) (see e.g. [16]).

Definition (7) requires that for all $\mathbf{\Gamma}$ where $f_0(\mathbf{\Gamma}) \neq 0$, $f_0\left(M^T S^t \mathbf{\Gamma}\right) \neq 0$, a condition referred to as *ergodic consistency*. Ergodic consistency guarantees that the probability *density* at one point of a Loschmidt's trajectory/antitrajectory conjugate pair is positive if the corresponding point has positive probability density (*i.e.* Loschmidt trajectory/antitrajectories pairs always exist).

When f_0 is the equilibrium distribution corresponding to no driving, Ω represents the energy dissipation, a quantity that can be computed or measured in experimental systems, regardless of how near or far the system is from equilibrium [16-23]. If S^t and f_0 are time reversal invariant (TRI), *i.e.* $M^T S^t = S^{-t} M^T$ and $f_0\left(M^T\mathbf{\Gamma}\right) = f_0(\mathbf{\Gamma})$, as appropriate for equilibrium probability densities, Ω is odd with respect to time reversal, $\Omega\left(M^T\mathbf{\Gamma}\right) = -\Omega(\mathbf{\Gamma})$, as appropriate for dissipation. Consequently, its average with respect to any TRI distribution f vanishes: $\langle\Omega\rangle_f = 0$. We make use of this property to introduce the condition called *ΩT-mixing*, *i.e.*:

$$\lim_{t\to\infty} \int_0^t \left\langle \left(A \circ S^s\right)\Omega \right\rangle_0 \mathrm{d}s = L_A \in \mathbf{R}. \tag{9}$$

Because $\langle\Omega\rangle_0 = 0$, $C_0(t) = \langle (A \circ S^t)\Omega \rangle_0$ and the correlation function is required by (9) to vanish faster than $1/t$. This condition is particularly useful in connection with the Dissipation Theorem [24] for the response of a phase variable *A*, which in general terms reads:

$$\langle A \rangle_t = \langle A \circ S^t \rangle_0 = \langle A \rangle_0 + \int_0^t \langle (A \circ S^s) \cdot \Omega \rangle_0 \mathrm{d}s \ . \tag{10}$$

III.1 An example: the isokinetic particle system

The dissipation function and the ΩT-mixing condition introduced above are especially useful in the framework of NEMD models, where Ω can then be identified with the energy dissipation rate divided by the instantaneous thermodynamics temperature of the underlying equation state the system would relax to if it was so allowed.

For concreteness, let us discuss these issues in the context of a system of *N* particles subject to the following equations of motion:

$$\dot{\mathbf{q}}_i = \mathbf{p}_i / m + C_i \mathbf{F}_e, \quad \dot{\mathbf{p}}_i = \mathbf{F}_i + D_i \mathbf{F}_e - S_i \alpha_{IK} \mathbf{p}_i + S_i \mathbf{F}_{th} \tag{11}$$

where $\mathbf{F}_e$ is an external dissipative field (*e.g.* an electric field applied to a molten salt), and the scalars C_i and D_i couple the system to $\mathbf{F}_e$. Let the N_{th} thermostatting particles belong to the set *th*, let S_i be a switch to determine whether particle *i* is a member of the set ($S_i = 0, i \notin th$, $S_i = 1, i \in th$), let the thermostat multiplier [7] α_{IK} be chosen to fix the kinetic energy of the thermostatting particles at the value K_{th}, and $\mathbf{F}_{th}$ be a fluctuating force fixing the momentum of the thermostatting particles, which is selected to have $\mathbf{P}_{th} = \sum_{i=1}^{N} S_i \mathbf{p}_i = \mathbf{0}$. We assume the interatomic forces $\mathbf{F}_i$, $i = 1,...,N$, are smooth and short ranged functions of the interparticle separation. We also assume

that in the absence of the thermostatting and momentum zeroing forces, the equations of motion preserve phase space volumes (*i.e.* $\partial/\partial\mathbf{\Gamma}\bullet\dot{\mathbf{\Gamma}}^{ad}\equiv\Lambda^{ad}(\mathbf{\Gamma})=0$) where $\mathbf{\Gamma}\equiv(\mathbf{q}_1,...,\mathbf{p}_N)$ is the phase space vector and *ad*, an abbreviation for adiabatic, means that the time derivative is calculated with thermostatting and momentum zeroing forces turned off. This condition is known as the adiabatic incompressibility of phase space, or AI$\mathbf{\Gamma}$ [7].

We assume the system of particles is subject to infinite checkerboard boundary conditions [7] – at least in the direction of the force. This means that angular momentum is not a constant of the motion. It also means that dissipation can go on forever without the system relaxing to equilibrium. Currents can flow in the direction of the force forever. In our system the application of infinite checkerboard boundary conditions means that space is translationally homogeneous but orientationally anisotropic. There are no walls with normals parallel to the field to stop particle currents. The thermostatting particles may be taken to be solid particles, like the walls parallel to the field, which can absorb or liberate heat that may be required to generate a NESS characterized by a fixed value for the kinetic energy of the thermostatting particles.

As the initial equilibrium distribution, we select the distribution that is invariant for the system (11) with vanishing $\mathbf{F}_e$. This is referred to as the isokinetic canonical distribution:

$$f_0(\mathbf{\Gamma})=\frac{\exp\left[-\beta_{th}H_0(\mathbf{\Gamma})\right]\delta(\mathbf{P}_{th})\delta\left(K_{th}(\mathbf{\Gamma})-K_{\beta,th}\right)}{\int\exp\left[-\beta_{th}H_0(\mathbf{\Gamma})\right]\delta(\mathbf{P}_{th})\delta\left(K_{th}(\mathbf{\Gamma})-K_{\beta,th}\right)\mathrm{d}\mathbf{\Gamma}} \tag{12}$$

where $K_{th}(\mathbf{\Gamma})=\sum S_i p_i^2/2m_i$ is the kinetic energy of the thermostatting particles and $K_{\beta,th}=(3N_{th}-4)/(2\beta_{th})$ is the fixed *value* of the kinetic energy of the thermostatting

particles. The number of particles in a unit cell is N. The kinetic energy of the thermostatting particles is fixed using the Gaussian multiplier α_{IK},

$$\alpha_{IK} = \frac{\sum_i S_i(\mathbf{F}_i + D_i\mathbf{F}_e + S_i\mathbf{F}_{th})\cdot\mathbf{p}_i}{\sum_i S_i\mathbf{p}_i\cdot\mathbf{p}_i}, \tag{13}$$

in the equations of motion. Here $\beta_{th} = 1/k_B T_{th}$ where k_B is Boltzmann's constant and for isokinetic systems T_{th} is the kinetic temperature of the thermostatting particles. It is also the equilibrium thermodynamic temperature the system will relax to if it is so allowed. Because the total momentum of the system averages to zero, the equilibrium internal energy of the N-particles in the unit cell is the average of $H_0(\mathbf{\Gamma}) = K(p) + \Phi(q)$ over the distribution f_0: $\langle H_0 \rangle_0 = \int H_0(\mathbf{\Gamma}) f_0(\mathbf{\Gamma}) \mathrm{d}\mathbf{\Gamma}$, where $K(p)$ and $\Phi(q)$ are respectively the kinetic and potential energy of all the particles in the original unit cell. For any particle in the original unit cell and at any time, the potential energy may involve interactions with particles that were not, or are not, located in the original unit cell.

We should now specify the ostensible phase space domain that is not referred to explicitly in Eq. (12). In the full canonical ensemble the particle momenta are unbounded, however the delta functions in the isokinetic canonical ensemble place four constraints on the momenta of some of the particles in the system so this is no longer the case. The initial coordinates of the particles will each be within some finite range, $\pm L$, within the unit cell of the periodic system. Due to the periodicity, any particle and its environment is essentially identical to any periodic image of that particle. Particles can always be "re-imaged" back into the original unit cell. However calculating certain quantities may have spurious discontinuities if this is done. Thermodynamic quantities like pressure, internal energy etc. are all continuous in

time, independent of whether particles are "imaged" in the unit cell. The thermostatting region that is unnatural can be made arbitrarily remote from the natural system of interest, so that it does not affect the bulk behaviour, cf. Section VI.

A key point in the definition (7) of the dissipation function, is that $\mathbf{\Gamma}$ and $M^T S^t \mathbf{\Gamma}$ are the initial phase points for a trajectory and its conjugate (antitrajectory) respectively. This places constraints on the propagator, S^t. For a system defined by Eq. (11), satisfying AI$\mathbf{\Gamma}$ and that is initially in equilibrium with distribution function (12), it is easy to show that Ω can be written as:

$$\Omega(\Gamma) = -\beta_{th} \mathbf{J}(\mathbf{\Gamma}) V \cdot \mathbf{F}_e \tag{14}$$

where V is the volume of the unit cell of our infinitely periodic system, the dissipative flux [7] is given by:

$$\sum_{i=1}^{N} [\frac{\mathbf{p}_i}{m} D_i - \mathbf{F}_i C_i] \bullet \mathbf{F_e} \equiv -\mathbf{J}(\mathbf{\Gamma}) V \bullet \mathbf{F_e} = \dot{H}_0^{ad} . \tag{15}$$

where $\dot{H}_0^{ad}$ is the rate of change of H_0 according to the adiabatic (unthermostatted equations of motion). Indeed, simple algebra yields [23, 26]:

$$\Lambda(\mathbf{\Gamma}) = \beta_{th} \left[\dot{H}_0(\mathbf{\Gamma}) + \mathbf{J}(\mathbf{\Gamma}) V \cdot \mathbf{F}_e \right] \tag{16}$$

and

$$\frac{f_0(\Gamma)}{f_0\left(\mathrm{M}^T S^t \Gamma\right)} = \exp\left\{ \beta_{th} \int_0^t \dot{H}_0\left(S^s \Gamma\right) \mathrm{d}s \right\} \tag{17}$$

In fact (14, 15) define what we call the *primary* dissipation function for this system. If the field is set to zero there is no dissipation because the initial distribution is the equilibrium distribution for the zero field dynamics (11). In the linear regime, the average dissipation function is equal to the so-called entropy production rate.

In the case of the equations of motion (11) and initial distribution (12), the Dissipation Theorem (10) can be written as:

$$\langle A\rangle_{t,\mathbf{F}_e} = \langle A\rangle_{0,0} - \beta_{th} V \int_0^t \left\langle \left(A \circ S^s\right)\mathbf{J}\right\rangle_{0,\mathbf{F}_e} \cdot \mathbf{F}_e \, \mathrm{d}s \tag{18}$$

where the various physical ingredients are explicated, and the notation stresses their roles: $\langle\cdot\rangle_{t,\mathbf{F}_e}$ is the ensemble average with respect to the phase space density f_t which evolves from the initial f_0 according to the full field-dependent dynamics, denoted for simplicity by S^t instead of $S^t_{\mathbf{F}_e}$; and therefore $\left\langle \left(A \circ S^s\right)\mathbf{J}\right\rangle_{0,\mathbf{F}_e}$ means the average with respect to f_0, with the evolution of *A* carried out with the field-driven dynamics.

Expressions (10, 18) are exact, arbitrarily near or far from equilibrium and also for systems of arbitrary size. They look similar, but they differ from the linear response expressions for the evolution of phase variables in that the time correlation functions are those determined with the field driven dynamics in (10, 18), whereas the equilibrium time correlation functions appear in linear response theory expressions. Equation (18) shows that if the driving field vanishes, the ensemble averages of phase functions are time independent, provided f_0 is invariant for the field-free dynamics. If the system starts with the equilibrium distribution (12), the distribution is preserved by the field free, thermostatted dynamics.

Using the definition (9), the average long time response of *A* given by (18) yields a real number in the long time limit,

$$\lim_{t\to\infty} \langle A\rangle_{t,\mathbf{F}_e} = \langle A\rangle_{0,0} + L_{A,\mathbf{F}_e} \tag{19}$$

if the system is ΩT-mixing. Property (9) is *necessary* and *sufficient* for this result, so another statement of ΩT-mixing in the present case is that all phase variables satisfy Eq. (19). This result is completely general, because Eq. (10) and its various versions are exact and directly derived from the dynamics. Moreover, Eq. (19) affords one *"a posterior"* test to assess whether the given dynamical system is ΩT-mixing or not.

Indeed, while the ΩT-mixing condition may either to hold or not, depending on the case at hand, analogously to ergodic properties of generic dynamical systems, ΩT-mixing may not be easy to assess *a priori*, from mere knowledge of the equations of motion, although it is weaker than ergodicity (see Ref. [25] for the case of simple low dimensional maps).

Current NEMD simulations concerning the validity of the Steady State Fluctuation Relation for Ωindicate that for standard choices of parameters, the isokinetic model treated in this section is indeed ΩT-mixing [26]. On similar grounds, we expect other Gaussian or Nosé-Hoover thermostatted systems in the fluid state to be ΩT-mixing.

IV. CONVERGENCE OF ENSEMBLE AVERAGES UNDER ΩT-MIXING

We aim to find the conditions under which convergence of ensemble averages also corresponds to relaxation to a NESS for a single system, sampled at random in the whole phase space. Let us begin by observing that the NESS attractors for dissipative dynamics concentrate probability on sets whose dimension is less than that of the ostensible phase space, and that this dimension decreases as the average dissipation increases [27]. For some dynamics more than one NESS attractor will exist. Therefore, different systems starting from different phases $\mathbf{\Gamma} \in \mathcal{M}$ could evolve towards different asymptotic states yielding different time averages, although they all started in the same initial equilibrium state, characterized *e.g.* by the initial distribution μ_0. Because the phase space averages (10, 18, 19) run over all initial phases, $\lim_{t\to\infty}\langle A\rangle_{t,\mathbf{F}_e}$ would then be a weighted average of the different averages pertaining to the different asymptotic states and, as such, it would not necessarily represent the results of any single experimental measurement. The problem cannot always be cured by separately considering the different basins of attraction in $\mathcal{M}$ because, in general they are too finely intertwined with each other and cannot be separated.

Let us consider the characteristic function χ_a^A of the invariant set[6] $E_a^A = \{\mathbf{\Gamma} \in \mathcal{M} : \overline{A}(\mathbf{\Gamma}) = a\}$, *i.e.* the set of phases $\mathbf{\Gamma}$ such that

$$\overline{A}(\mathbf{\Gamma}) = \lim_{t\to\infty} \frac{1}{t}\int_0^t A(S^s\mathbf{\Gamma})\,\mathrm{d}s = a \tag{20}$$

This set is invariant because of the limit in *t*, and is disjoint from any other E_b^A with $b \neq a$ because the time average is uniquely determined by Γ. Because time averages

[6] A set E in $E \subset \mathcal{M}$ is called an invariant set for the dynamics if $E = S^t E = S^{-t} E$.

exist quite generally [28], let us assume that the limit in equation (20) exists for all phases except a set of vanishing phase space volume. Then, the union over all a, $\bigcup_{a\in\mathbf{R}} E_a^A$, constitutes a set of μ_0-measure 1: $\mu_0\left(\bigcup_{a\in\mathbf{R}} E_a^A\right)=1$.

It may happen that $\mu_0\left(E_{\hat{a}}^A\right)=1$ for a given value $\hat{a}$. This means that the time averages along single trajectories (single experimental determinations of the property A) equal $\hat{a}$, apart from trajectories whose initial conditions constitute a set of vanishing μ_0 measure. In this case, the ensemble average computed with μ_0 equals the time average $\hat{a}$, because points leading to time averages other than $\hat{a}$ have vanishing weight in the ensemble average. In general, this does not need to be the case and, as usual in response theory, ΩT-mixing only constitutes a condition for relaxation on average, rather than for single experiments.

We conclude this section observing that the characteristic function of invariant sets, such as the unions of sets like E_a^A, obey the ΩT-mixing condition, as it can be directly gathered from the definitions. Indeed, for every finite time t, Eq. (10) and the equality $\mu_t(E)=\mu_0\left(S^{-t}E\right)=\mu_0(E)$, yield:

$$\mu_t\left(E_a^A\right)=\left\langle\chi_a^A\right\rangle_t=\left\langle\chi_a^A\right\rangle_0+\int_0^t\left\langle\left(\chi_a^A\circ S^s\right)\Omega\right\rangle_0\mathrm{d}s=\mu_0\left(E_a^A\right)=\left\langle\chi_a^A\right\rangle_0 \tag{21}$$

which means

$$\left\langle\left(\chi_a^A\circ S^t\right)\Omega\right\rangle_0=\int\chi_a^A(S^t\boldsymbol{\Gamma})\Omega(\boldsymbol{\Gamma})f_0(\boldsymbol{\Gamma})\mathrm{d}\boldsymbol{\Gamma}=\int_{E_a^A}\Omega(\boldsymbol{\Gamma})f_0(\boldsymbol{\Gamma})\mathrm{d}\boldsymbol{\Gamma}=0 \tag{22}$$

Furthermore, any constant of the motion will also trivially satisfy Eq. (19).

V. WEAK T-MIXING

To tackle the problem of tipicality for relaxation to nonequilibrium steady states, we adopt the perspective developed in the study of the steady state fluctuation relations [23], which pointed out a dynamical property, later called wT-mixing. Let f_0 be an initial probability density in the phase space $\mathcal{M}$. The dynamics S^t is called wT-mixing with respect to $\mathrm{d}\mu_0 = f_0 \mathrm{d}\mathbf{\Gamma}$ if

$$\left[\left\langle \left(A \circ S^t\right) B\right\rangle_0 - \left\langle A \circ S^t\right\rangle_0 \langle B\rangle_0\right] = \left[\left\langle \left(A \circ S^t\right) B\right\rangle_0 - \langle A\rangle_t \langle B\rangle_0\right] \equiv C_0(t) \xrightarrow[t\to\infty]{} 0 \tag{23}$$

where we used the identity $\left\langle A \circ S^t\right\rangle_0 = \langle A\rangle_t$.[7] The second equality of Eq. (23) introduces the correlation function $C_0(t)$ of A and B, with respect to the initial distribution f_0; unlike mixing, its decay in time is due to the loss of correlations between the initial and the evolving probability distributions, $\mathrm{d}\mu_0 = f_0 \mathrm{d}\mathbf{\Gamma}$ and $\mathrm{d}\mu_t = f_t \mathrm{d}\mathbf{\Gamma}$.

V.1 Strong ensemble relaxation

Suppose that wT-mixing holds for two observables A and $B = h_0 / f_0$, where h_0 is a smooth positive function that vanishes outside a given hypersphere $E \subset \mathcal{M}$ (which is not necessarily an invariant set) of positive radius. Without loss of generality, we may take $\int h_0(\mathbf{\Gamma})\mathrm{d}\mathbf{\Gamma} = 1$, which makes h_0 a probability density supported on E. The validity of Eq. (23) yields:

[7] The dynamics is called T-mixing if $\lim_{t\to\infty} \int_0^t \left[\left\langle \left(A \circ S^s\right) B\right\rangle_0 - \langle A\rangle_t \langle B\rangle_0\right] ds \equiv \lim_{t\to\infty} \int_0^t C_0(s)\, ds = L_A \in \mathbf{R}$, which will be the case if the correlation function decays more quickly than $1/t$. wT-mixing is therefore a weaker condition than T-mixing that does not specify the rate of convergence.

$$\left[\int f_0(\mathbf{\Gamma})B(\mathbf{\Gamma})A\left(S^t\mathbf{\Gamma}\right)\mathrm{d}\mathbf{\Gamma}-\int f_0(\mathbf{\Gamma})A\left(S^t\mathbf{\Gamma}\right)\mathrm{d}\mathbf{\Gamma}\int f_0(\mathbf{\Gamma})B(\mathbf{\Gamma})\mathrm{d}\mathbf{\Gamma}\right]$$

$$=\left[\int h_0(\mathbf{\Gamma})A\left(S^t\mathbf{\Gamma}\right)\mathrm{d}\mathbf{\Gamma}-\int f_0(\mathbf{\Gamma})A\left(S^t\mathbf{\Gamma}\right)\mathrm{d}\mathbf{\Gamma}\right]=\left[\langle A\rangle_t^{(h)}-\langle A\rangle_t^{(f)}\right]\xrightarrow[t\to\infty]{}0 \tag{24}$$

where $\langle A\rangle_t^{(h)}$ and $\langle A\rangle_t^{(f)}$ are the phase space averages of A at time t, starting from the ensembles h_0 and f_0, respectively. Equation (24) states that provided the asymptotic observable value $\langle A\rangle_\mu$ exists (something guaranteed under ΩT-mixing), its value does not depend on the initial ensemble.[8] As a matter of fact, because f_0 is an equilibrium distribution, we can assume that it is approximately constant within phase space hyperspheres of sufficiently small radius. If the probability densities supported over any hypersphere verify wT-mixing with A, the ensemble averages of A over trajectories starting within all hyperspheres yield the same value. This does not mean that almost all single system time averages converge to that value, as every hypersphere could contain a positive fraction of initial conditions of trajectories producing different time averages. However, it requires the fractions of initial conditions leading to different time averages to be the same everywhere in $\mathcal{M}$, something peculiar in relation to the phase space description of a physical object and to the measurements that identify its macroscopic state. While, mixing refers to invariant measures, wT-mixing refers to the known initial distributions. This has many advantages, including that for our result to hold, it does not matter whether a single steady state attractor is approached in time or not.

V.2 Single system relaxation

If A and B are the characteristic function χ_E of a set E, Eq. (23) implies:

$$\left[\mu_0(S^{-t}E\cap E)-\mu_0(S^{-t}E)\mu_0(E)\right]=\left[\mu_t(E\cap S^tE)-\mu_t(E)\mu_0(E)\right]\xrightarrow[t\to\infty]{}0 \tag{25}$$

[8] Note: We assume $f_0>0$ on all $\mathcal{M}$, while h_0 does not need to be so.

If E is also invariant, its probability does not change in time so $\mu_t(E \cap S^t E) = \mu_t(E) = \mu_0(E)$, and Eq. (25) shows that $\mu_0(E) = \mu_0(E)^2$, *i.e.* $\mu_0(E) = 0$ or 1. Clearly this would be inconsistent with the existence of more than one disjoint invariant set of non-zero μ_0-measure. Therefore wT-mixing can be taken to imply that there is only one such set. This uniqueness argument cannot be extended to the attractors of dissipative systems, because they have vanishing volume.

If the initial distribution was taken as the invariant steady state distribution, (25) would be satisfied by ergodic measures, but even in this case wT-mixing affords a fresh perspective on ergodic notions in physics, because it has been developed to reconcile aspects of physical systems, such as the validity of fluctuation relations, with dynamical systems theory, cf. [22, 23, 26].

More precisely, the only formal difference between Eq. (4) and Eq. (23) is that Eq. (4) refers to an invariant measure, while the initial distribution of Eq. (23) is not invariant under nonequilibrium dynamics. This leads to a major conceptual difference between the notions of mixing and wT-mixing. While mixing strictly speaks only of the decay of correlations between events within a given steady state, wT-mixing speaks of the loss of correlations between the initial and the evolving probability distributions. In cases in which these distributions characterize the macroscopic states of a given object, this affords a description of the relaxation process.

Then, let us consider a system where E_b^A is the set of all points in phase space with $\overline{A}(\mathbf{\Gamma}) = b$. As discussed below Eq. (20), this is an invariant set. Suppose that condition (24) holds for sets such as $E_{a,\delta}^A = \bigcup_{b \in [a, a+\delta)} E_b^A$, for $\delta > 0$, that are invariant because they are a union of invariant sets. Therefore, given two different

values *a* and *b*, Eq. (24) implies that $\mu_0(E^A_{a,\delta})$ and $\mu_0(E^A_{b,\delta})$ equal 0 or 1, and one of them at most can be 1, if $|b-a|>\delta$. In addition, for any $\delta>0$, the real numbers $\mathbf{R}$ can be expressed as a countable union of disjoint sets like $[a,a+\delta)$:

$$\mathbf{R}=\bigcup_{n\in\mathbf{Z}}[n\delta,(n+1)\delta) \quad \text{and} \quad \mu_0\left(\bigcup_{n\in\mathbf{Z}}E_{[n\delta,(n+1)\delta)}\right)=1\,. \tag{26}$$

Whatever accuracy $\delta>0$ we choose, one of the invariant sets must have μ_0 -measure one and all the rest must be of μ_0 -measure zero. Then, if the system is wT-mixing, for any $\delta>0$ there is a single $a\in\mathbf{R}$ such that $\mu_0(E^A_{a,\delta})=1$, and $\mu_0(E^A_{b,\delta})=0$ for all *b*, with $|b-a|>\delta$. In other words, wT-mixing for the invariant sets $E^A_{a,\delta}$ implies that all single systems, apart from a set of vanishing phase space volume, converge to a state in which the measurements of *A* yield, with arbitrary accuracy, the value *a*.

This does not imply that there is a single NESS attractor: a system with various NESS attractors each with $a\le\overline{A}<a+\delta$, can be wT-mixing if there are no other sets with non-zero μ_0 -measure for which $\overline{A}<a$ or $\overline{A}\ge a+\delta$. The intricacies of the phase space structure are irrelevant in our approach, as they should be, because what matters physically are the values of the observables. Because the physically relevant observables required to characterise a physical system are but a few, condition (25) does not appear particularly strong, if the analysis is restricted to them.

The result is not merely an "average" relaxation concerning an ensemble of systems (as usual in response theory), but it describes relaxation as expected for single thermodynamic systems, and points out a new kind of typicality for dissipative systems: all (but those in a negligible volume of phase space) relax to the same steady state, as far as measurements of observables can tell. We use the term *physical ergodicity* to refer to the condition in which time averages, or physical measurements,

of a given observable yield a unique value, except for a set of vanishing Lebesgue measure, even in the case of dissipative dynamics.

Note that this situation is not the one of standard ergodic theory, because it concerns the initial distribution of phases and not the steady state. This is important, since "*almost all*" points on the NESS attractor of dissipative systems might satisfy the equality of time-averages and ensemble averages over the points on the NESS attractor, however these points have zero μ_0-measure and therefore refer to *"almost none"* in terms of phase space volumes. Moreover, our uniqueness of the time averages, does not need the uniqueness of the attractor, hence it is weaker than metric transitivity. wT-mixing for the invariant sets $E^A_{a,\delta}$ represents a condition that the dynamics must obey to ensure relaxation from almost all initial phases. For Hamiltonian systems, it is not stronger than ergodicity, hence it is weaker than mixing. It can be graded without causing mathematical inconsistencies, by selecting the observables of interest, hence it can be made as weak as needed. Equation (25) for the invariant sets of the observables of interest is also trivially necessary: if $0 < \mu_0\left(E^A_{a,\delta}\right) < 1$, the condition is violated.

VI. PHYSICAL CONSIDERATIONS

Here we summarise the different transient mixing conditions and their physical implications. ΩT-mixing requires the correlation function with respect to the initial distribution not only to go to zero, but (whether stationary or transient) to vanish sufficiently rapidly for its integral to converge, so that $\lim_{t\to\infty}\langle A\rangle_t$ exists and is *finite*. For example if the equilibrium time correlation function goes as $1/t$ at long times, we will have a logarithmic divergence and the system will not relax to a NESS. This is quite different to the ergodic theory result for autonomous Hamiltonian systems, where mixing implies relaxation on average towards the time independent microcanonical equilibrium distribution, irrespectively of the decay rate of the correlations. However, this is due to the fact that mixing concerns states with measures that do not evolve. The mixing condition (4-6) cannot be used to prove relaxation from a smooth initial distribution to the invariant NESS distribution because the invariant nonequilibrium distribution is singular [15].

If we turn briefly to the *transient* time correlation function for the nonlinear response, the mixing condition is simply not relevant. The distributions of states used to compute transient time correlation functions are not stationary.

Equation (18) can be used to derive the Green-Kubo [20, 29, 30] relations in the limit of zero field. However the conditions required are subtle, and different to that used to obtain (18). Kubo's results [29] were for the linearized adiabatic response (*i.e.* no thermostats) of a canonical ensemble of systems. We derived Eq. (18) for isokinetic dynamics where the kinetic energy of the thermostatting particles is fixed and the distribution for the system of interest is isokinetic canonical – Eq. (12). Thus the equilibrium time correlation function appearing in (18) is for field free isokinetic dynamics. Therefore to obtain the Green-Kubo relationships, we need to derive the

equivalent of (18) using an initial canonical distribution function to generate the initial points, followed by unthermostatted equations of motion to evaluate the correlation function. The resulting equation will look like (18), but as in Kubo's system the time correlation functions will involve canonical distributions but field free, constant energy, Newtonian trajectories. We also note that Eq. (18) using the isokinetic dynamics and starting from an isokinetic distribution for ΩT-mixing systems are consistent with the result of Evans and Morriss [31] where it was proved that to leading order in the number of degrees of freedom in the system with a correction of order *O(1/N)*, the two equilibrium correlation functions are identical. Of course if the dissipative field only couples to particles in the system of interest and the thermostat region is large and remote, the fluctuations in the dissipation function (which is local to the system of interest) will hardly be affected by the presence or absence of thermostatting terms in the large remote thermostatting region. Because the thermostat is unphysical, we can make the system more realistic by only thermostatting a remote set of particles. If the external fields are set to zero and the system is allowed to relax to equilibrium we know the thermodynamic temperature of that underlying equilibrium system. That is the temperature that appears in the equations given above.

There is yet another interesting observation we can make regarding Kubo's system [29, 30]. If you consider viscous flow in a dilute gas then as is known from kinetic theory, the viscosity of a gas increases with temperature. This means that for any finite field, no matter how small, the shear stress of an adiabatic shearing gas must increase with time. This means that a shearing unthermostatted gas can never be ΩT-mixing! In a physical sense for such a system, time correlations either do not decay or do not decay rapidly enough for ΩT-mixing. One can see how this memory

effect occurs. If among the initial ensemble members, one encounters a fluctuation that increases the gas viscosity, that fluctuation will, at fixed strain rate, heat the gas slightly. In this slightly heated gas the viscosity will be slightly higher than on average increasing the likelihood of further fluctuations that in turn increase the viscosity. This is a run-away process that prevents the decay of correlations required for the ΩT-mixing condition.

If we assume ΩT-mixing we see that although the long time states predicted by (10, 18, 19) may not be ergodic in the metrically transitive sense, those asymptotic states have nevertheless stationary ensemble averages. If wT-mixing holds for the characteristic functions of the invariant sets of the different values of the observables of interest, the ensemble averages equal the corresponding single system time averages, for almost all initial phases: exceptions constitute a set of zero volume.

VI. CONCLUSION

We have shown that ΩT-mixing is necessary and sufficient for an initial ensemble to relax to a steady state. wT-mixing, instead, leads to convergence in the sense of time averages of observables, from almost all initial conditions, *i.e.* for practically all single experimental measurements of a system. We refer to this situation as *Physical Ergodicity*, because it preserves the notion of typicality that we consider more appropriate for physics, even in case of dissipative dynamics: typicality in the sense of the Lebesgue measure.

These conditions differ substantially from the standard ergodic theory notions, because they refer to the initial probability distributions, and not to invariant measures. Among the numerous consequences of this fact, we have that the relaxation argument expressed by Eq. (4), makes no reference to the rate at which correlations decay in contrast to that based on ΩT-mixing. This reflects the fact that the argument of Eq. (4) speaks of the decay of correlations within the NESS attractor, and not of the decay of correlations between an initial and final state.

The condition based on wT-mixing better suits the needs of physical studies of nonequilibrium many-body interacting particles than arguments based on mixing within the NESS attractor because the NESS attractor of this dissipative dynamics occupies zero volume in the ostensible phase space. Therefore even if "almost all" points on the NESS attractor satisfy a given desired property, these points have zero measure in the equilibrium distribution and therefore this means very little for systems starting in a given equilibrium state. Moreover, the wT-mixing condition can be graded to the needs of observations, by restricting it to the variables of physical interest. This frees the dynamics of demanding conditions such as metric transitivity.

If the system is wT-mixing, there may be a set of initial conditions of vanishing

μ_0-volume measure that do not yield the same time average, but this is only a set of vanishing phase space volume.

In other words, even in the case of dissipative dynamics the irreversibility of the relaxation process has been connected to "counting" of states, as done in the past for the equilibrium case. In that case, "counting" meant comparing the fractions of phase space pertaining to different states, *e.g.* Ref. [32], and finding that by far the largest fraction is occupied by the equilibrium state [8, 33-35]. In the case of convergence to a NESS of a wT-mixing system, we have shown that by far the largest fraction of the phase space is occupied by phases that yield the same observable value for a given phase variable.

wT-mixing for smooth functions implies a weaker result related to relaxation than full wT-mixing does. This is, however, a rather strong result that states that relaxation on average does not depend on the initial distribution, if this limited wT-mixing holds. In particular, although this is just an ensemble result, the initial ensemble can be as small a set around any $\Gamma \in \mathcal{M}$ as one wishes.

A system that relaxes to a steady state is OmegaT-mixing. If the system relaxes from all initial conditions, apart from a set of vanishing volume, to give the same time-independent properties , then it might also be wT-mixing. Mathematically assessing whether a given model verifies our transient mixing conditions may not be easy in general, and so far only some simple maps have been treated explicitly in these terms [25]. However, numerical evidence suggests that NEMD models of simple fluids can be ΩT-mixing, and could be wT-mixing as well [26].

ACKOWLEDGEMENTS

We thank Ms. Charolotte F. Petersen, Dr Owen G. Jepps and Dr James C. Reid for useful comments. We would also like to thank the Australian Research Council for support of this research. LR thanks the European Research Council, for funding under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n 202680. The EC is not liable for any use that can be made on the information contained herein.

REFERENCES

1. G. Gallavotti, *Ergodicity, ensembles, irreversibility in Boltzmann and beyond*, J. Stat. Phys. **78**, 1571 (1994)
2. J. Uffink, *Boltzmann's Work in Statistical Physics*, Stanford Encyclopedia of Philosophy (2014)
3. E. Fermi, *Beweis, daß ein mechanisches normalsystem im algemeinen quasi-ergodisch ist*. Physikalische Zeitschrift **24**, 261 (1923). Reprinted in Fermi, E.: *Note e Memorie (Collected papers)*, Accademia dei Lincei and University of Chicago Press, vol. I (1961)
4. A. Khinchin, *Mathematical Foundations of Statistical Mechanics*, Dover Publications, New York, (1949)
5. P. Mazur, J. van der Linden, *Asymptotic form of the structure function for real systems*, J. Math. Phys. **4**, 271 (1963)
6. L. Cerino, F. Cecconi, M. Cencini, A. Vulpiani, The role of the number of degrees of freedom and chaos in macroscopic irreversibility, Physica A **442**, 486-497 (2016)
7. D.J. Evans, G.P. Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, Cambridge University Press, Cambridge (2008)
8. P. Castiglione, M. Falcioni, A. Lesne, A. Vulpiani, *Chaos and Coarse Graining in Statistical Mechanics*, Cambridge University Press, Cambridge (2008)
9. G. Gallavotti (ed), *The Fermi-Pasta-Ulam Problem: A Status Report*, Springer, Berlin (2008)
10. P. Poggi and S. Ruffo, *Exact solutions in the FPU oscillator chain*, Physica D **103**, 251-272 (1997)
11. J. Milnor, *On the Concept of an Attractor*, Comm. Math. Phys. **99**, 177 (1985)

12. I.P. Cornfeld, S.V. Fomin, Ya. G. Sinai, *Ergodic Theory*, Springer-Verlag, New York (1982)
13. Ya. G. Sinai, *Introduction to Ergodic Theory*, Princeton University Press, Princeton (1976)
14. 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)
15. D.J. Evans, S.R. Williams, D.J. Searles and L.Rondoni, *On the relaxation to nonequilibrium steady states,* arXiv:1602.05808 [cond-mat.stat-mech] (2016).
16. D.J. Evans, D.J. Searles and S.R. Williams, *Dissipation and the relaxation to equilibrium*, J. Stat. Mech. P07029 (2009)
17. J.C. Reid, D.J. Evans and D.J. Searles, *Communication: Beyond Boltzmann's H-theorem: Demonstration of the relaxation theorem for a non-monotonic approach to equilibrium*, J. Chem. Phys. **136**, 021101 (2012)
18. D.J. Evans, D.J. Searles and S.R. Williams, *On the probability of violations of Fourier's law for heat flow in small systems observed for short times*, J. Chem. Phys. **132**, 024501 (2010)
19. D.J. Searles and D.J. Evans, *Ensemble dependence of the transient fluctuation theorem*, J. Chem. Phys. **113** 3503 (2000)
20. D.J. Evans and D.J. Searles, *The fluctuation theorem*, Adv. Phys. **52**, 1529 (2002)
21. J.C. Reid, S.R. Williams, D.J. Searles, L. Rondoni and D.J. Evans, *Fluctuation Relations and the Foundations of Statistical Thermodynamics:a Deterministic approach and numerical demonstration*, in Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond, Editors R. Klages, W. Just, C. Jarzynski, Wiley-VCH, (2013)

22. D.J. Evans, D.J. Searles and L. Rondoni, *On the application of the Gallavotti-Cohen fluctuation relation to thermostatted steady states near equilibrium*, Phys. Rev. E **71**, 056120 (2005)

23. D.J. Searles, L. Rondoni, D.J. Evans, *The steady state fluctuation relation for the dissipation function*, J. Stat. Phys. **128**, 1337 (2007)

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

25. L. Rondoni, G. Dematteis, *Physical ergodicity and exact response relations for low-dimensional maps*, Comput. Meth. Sci. Tech. (to appear)

26. D. J. Searles, B. M. Johnston, D. J. Evans, L. Rondoni, Time Reversibility, Correlation Decay and the Steady State Fluctuation Relation for Dissipation, Entropy **15**, 1503 (2013)

27. D.J. Evans, E.G.D. Cohen, D.J. Searles, F. Bonetto, *Note on the Kaplan Yorke Dimension and Linear Transport Coefficients*, J. Stat. Phys, **101**, 17 (2000)

28. J. Palis, *A global perspective for non-conservative dynamics,* Ann. I. H. Poincaré **22**, 485 (2005)

29. R. Kubo, *Statistical-Mechanical Theory of Irreversible Processes. I. General Theory and Simple Applications to Magnetic and Conduction Problems*, J. Phys. Soc. Jpn. **12**, 570 (1957)

30. M.S. Green, *Markoff Random Processes and the Statistical Mechanics of Time - Dependent Phenomena. II. Irreversible Processes in Fluids*, J. Chem. Phys. **22**, 398 (1954)

31. D.J. Evans and G.P. Morriss, *Equilibrium time correlation functions under gaussian isothermal dynamics*, Chem. Phys. **87**, 451 (1984)

32. I. Pitowsky, *Typicality and the Role of the Lebesgue Measure in Statistical Mechanics*, in Y. Ben-Menahem and M. Hemmo (eds.), Probability in Physics, Springer, Berlin (2012)

33. S. Chibbaro, L. Rondoni, A. Vulpiani, *On the Foundations of Statistical Mechanics: Ergodicity, Many Degrees of Freedom and Inference*, Comm. Theor. Phys. **62**, 69 (2014)

34. S. Chibbaro, L. Rondoni, A. Vulpiani, *Reductionism, Emergence and Levels of Reality*, Springer, New York (2014)

35. J.L. Lebowitz, *Boltzmann's entropy and time arrow*, Physics Today **46**, 32 (1993)