POLITECNICO DI TORINO
Repository ISTITUZIONALE

Dissipation function: Nonequilibrium physics and dynamical systems

(Article begins on next page)

16 July 2022

Dissipation Function: Nonequilibrium Physics and Dynamical Systems

Salvatore Caruso [1], Claudio Giberti [2] and Lamberto Rondoni [3,4,*]

[1] Department of Physics, Informatics and Mathematics, Università di Modena and Reggio Emilia, via G.Campi 213/b, I-41125 Modena, Italy; 236561@studenti.unimore.it
[2] Department of Sciences and Methods for Engineering, Università di Modena and Reggio Emilia, via G.Amendola 2, I-42122 Reggio Emilia, Italy; claudio.giberti@unimore.it
[3] Department of Mathematical Sciences Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy
[4] INFN, Sezione di Torino, Via Pietro Giuria 1, 10125 Torino, Italy
* Correspondence: lamberto.rondoni@polito.it

Abstract: An exact response theory has recently been developed within the field of Nonequilibrium Molecular Dynamics. Its main ingredient is known as the Dissipation Function, Ω. This quantity determines nonequilbrium properties like thermodynamic potentials do with equilibrium states. In particular, Ω can be used to determine the exact response of particle systems obeying classical mechanical laws, subjected to perturbations of arbitrary size. Under certain conditions, it can also be used to express the response of a single system, in contrast to the standard response theory, which concerns ensembles of identical systems. The dimensions of Ω are those of a rate, hence Ω can be associated with the entropy production rate, provided local thermodynamic equilibrium holds. When this is not the case for a particle system, or generic dynamical systems are considered, Ω can equally be defined, and it yields formal, thermodynamic-like, relations. While such relations may have no physical content, they may still constitute interesting characterizations of the relevant dynamics. Moreover, such a formal approach turns physically relevant, because it allows a deeper analysis of Ω and of response theory than possible in case of fully fledged physical models. Here, we investigate the relation between linear and exact response, pointing out conditions for the validity of the response theory, as well as difficulties and opportunities for the physical interpretation of certain formal results.

Keywords: dissipative systems; entropy production; response theory; nonequilibrium steady states; nonequilibrium molecular dynamics; formal thermodynamics

1. Introduction

Since the 1980s, a dynamical systems approach to nonequilibrium phenomena has been developed, side by side with the increase of computing power and the ability to perform Molecular Dynamics (MD) simulations. Classical MD consists of algorithms meant to solve Newton's dynamical equations for interacting particle systems, that may be subjected to external fields. Such investigations result quite informative and practically useful, when quantum effects are irrelevant or negligible, which is the case of fluids under standard conditions.

Provided no dissipative forces are present, one speaks of equilibrium MD, which is widely used together with statistical mechanical relations, in order to compute: rheological properties, polymers behaviours in porous media, defects in crystals, friction between surfaces, formation of atomic clusters, features of biological macromolecules, formation of drugs, etc. The results are most satisfactory, and, in practice, difficulties only arise if the interatomic forces are too complicated to

properly implement, if the required number of simulated particles is too large, or the simulation times too long. In fact, the statistical relations connecting microscopic dynamical quantities to macroscopic properties provide quite an accurate and complete description of material objects. Therefore, MD is used to properly interpret results of experiments, and even in place of expensive or impossible experiments. For instance, MD may illustrate the formation of crystal defects, the propagation of fracture fronts inside solids, and the thermal dilation of nuclear fuel pellets inside a nuclear reactor, just to mention a few examples. The relevant literature is enormous, and dates back to early days of electronic computers, the Fermi–Pasta–Ulam–Tsingou problem being a forerunner that opened the way to an incredible number of mathematical, computational and physical advances [1,2]; see e.g., the books [3–7].

When the system of interest is driven away from equilibrium by external forces that continually feed energy in it and, at the same time, it is in contact with some environment that removes energy from it, a nonequilibrium steady state may be reached. The question is how to effectively model the typically very large environment [8]. Various methods have been developed for that. In particular, researchers introduced "synthetic forces" meant to efficiently represent the effect of a bath or reservoir on the property of interest for the system of interest, without simulating the atoms of the environment. Nonequilibrium MD (NEMD) was thus born, focusing on specific quantities of specific systems [6,9].

The term "synthetic" means that such forces do not exist in nature; they merely constitute a *"mathematical device used to transform a difficult boundary condition problem, the flow of heat in a system bounded by walls maintained at differing temperatures, into a much simpler mechanical problem"*, cf. the Introduction of Ref. [6]. In practice, the synthetic forces constrain in various ways the dynamics of particle systems fixing, for instance, the average of the fluctuating kinetic temperature. Such an approach is so successful that the corresponding algorithms are now provided as default routines of MD packages [10,11].

If the constraints are holonomic and the other forces conservative, the resulting equations of motion are Hamiltonian [12]; otherwise they are non-Hamiltonian and, in particular, they are dissipative, in the sense that on average they contract phase space volumes, although they remain time reversal invariant [6]. Consequently, from the angle of dynamical systems, their invariant measures are singular with respect to the Lebesgue measure. This approach may sound odd, if one expects all features of the dynamics to be derived from a Hamiltonian, but it simply replicates the usual construction of effective models, that bridge the gap between different levels of descriptions of natural phenomena, [13]. For instance, the Boltzmann equation and the Langevin equation alter the Hamiltonian dynamics to the point that it becomes irreversible and not deterministic; at the same time their success and foundational importance is not questionable. Justifications of the NEMD approach, within its specific realm, have been provided, demonstrating, for instance, that it properly computes transport coefficients [6]. Moreover, NEMD has been shown to provide a route to extend to dissipative dynamics the theory of equivalence of ensembles, as well as Khinchin's approach to ergodic theory in physics [6,9,14–16].

The point is the following: no mathematical model spans all space and time scales of interest for a given material object: a model is acceptable when it describes a certain set of properties that are observable within given space and time scales. Of course, the wider the range of applicability of the model, the better it is. In this respect, classical Hamiltonian dynamics is no exception. It constitutes a tremendously effective microscopic model for a vast variety of non-dissipative phenomena, but dissipation makes its use problematic, cf. recent investigations on kinetic theory [17,18]. Analogously, models derived within the framework of NEMD, such as Gaussian thermostatted dynamics, or Nosé–Hoover thermostatted dynamics can be accepted as valid models of certain phenomena. The question is how many aspects of a given phenomenon are properly described by NEMD models. As generally true in physics, interesting results and relations can be derived in uncharted territories, but only experience can eventually confirm or refute their validity.

As a matter of fact, NEMD has become popular for its effectiveness in computing transport coefficients, such as the viscosity of a fluid [6]. It was also realized that such coefficients can be obtained from dynamical systems properties such as the sum of Lyapunov exponents or, under special conditions, just from the sum of the largest and of the smallest exponent [6]. In turn, $-\Lambda$, the opposite of the phase space volumes variation rate Λ, which turns out to be proportional to the thermostatting (fictitious or synthetic) term may also be proportional to the entropy production σ when thermodynamics applies, see e.g., Ref. [19]. (A similar idea had previously been independently investigated by Andrej, who followed a different approach [20]). This idea became very popular when Evans, Cohen and Morriss applied a representation of SRB measures to the fluctuations of Λ in a Gaussian isoenergetic shearing system [21], as in that case such a quantity equals the entropy production, when local thermodynamic equilibrium (LTE) holds. That paper derived and tested the Fluctuation Relation (FR), one of the rare exact relations for nonequilibrium thermodynamics based on microscopic deterministic dynamics. For a time reversal invariant model of shearing fluids, the FR states that positive values of the energy dissipation, are exponentially more probable than their opposite. This was interpreted as an explanation of the second law of thermodynamics, and motivated a flurry of activity. Moreover, it led various authors to identify the entropy production with the phase space contraction [22,23].

Although Λ is relatively easy to handle in dynamical systems theory, it was pointed out that its identification with σ is only justified in a limited set of cases [24,25]. Furthermore, in time dependent settings and in presence of fluctuations, the connection of Λ with physically measurable properties results rather loose. It was then shown that the physically relevant quantity in the cases described by NEMD is the so called Dissipation Function Ω, which preserves the meaning of energy dissipation rate even when LTE is violated, and thermodynamic quantities such as σ do not exist. Only in a few cases Ω equals $-\Lambda$, although steady state averages of Ω and $-\Lambda$ may be simply related to each other.

Later Ω became the basis for a general response theory, in which it plays a role as fundamental for nonequilibrium states as thermodynamic potentials do for equilibrium states. It can be used to solve in terms of physically relevant quantities the Liouville equation and, consequently, to investigate the exact average response of the observables of ensembles of systems. Furthermore, Ω can be used to provide conditions for the evolution of single systems, [26,27]. (This should be contrasted with the fact that almost invariably response of systems to perturbations is given in terms of ensembles, and that a comparatively limited amount of works have tackled response beyond its linear approximations). Most importantly, these results are expressed in terms of a physically measurable quantity, rather than in terms of an abstract phase space quantity.

Unfortunately, it is rather complicated to thoroughly investigate the properties of the dissipation function in realistic particle models. Therefore, we consider mathematically amenable dynamical systems, even if their physical relevance may be limited, as often and successfully done in statistical physics, with the purpose of illustrating aspects of the relation between dynamics and thermodynamics; think for example of Helmoltz's celebrated theorem [28]. Recently, Qian, Wang and Yi have also followed this route, scrutinizing a number of interesting examples and investigating them in terms of the phase space contraction rate $-\Lambda$ [29]. The result is a set of intriguing relations between different dynamical quantities and formal thermodynamic quantities. In consideration of the fact that small systems, hence non-thermodynamic fluctuating quantities, as well as time dependent situations are ever more relevant in science and technology, we are going to perform a similar analysis within the dissipation function formalism.

This paper is organized as follows. Section 2 summarizes the theory of the Dissipation Function, also analyzing exactly solvable examples and numerical simulations. Section 3 compares linear Green–Kubo and exact response. Section 4 follows in part Ref. [29], using Ω rather than Λ. The two approaches appear as largely equivalent, since they both arise from the solution of the Liouville equation. Section 5 explicitly solves the cases of simple attractors, and of small dissipation, in the sense of small Λ. Section 6 discusses with examples the conditions under which the exact response

theory based on Ω holds, showing that Ω has to be continuous at least, and that the initial phase space probability density f_0 has to be positive in all the space that can be explored by the dynamics. This condition, called ergodic consistency [30] is reminiscent, but weaker than the requirement that initial and asymptotic measure be absolutely continuous with respect to each other, also called equivalence [31]. If that is not the case, modifications to the response formulae are required [32].

Finally, Section 7 summarises our work, highlighting two points in particular: investigating simple systems allows us to understand properties of Ω that are not immediately evident in realistic particle systems; on the other hand we propose a kind of thermodynamic formalism for the analysis of dynamical systems, cf. Ref. [33], that is based on the language of nonequilibrium, rather than of equilibrium, statistical mechanics. Section 7 also stresses the role of Ω as the energy dissipation, in the case of interacting particle systems.

Remark: one of the anonymous referees pointed out that the scope of our investigation can be substantially extended, because: *"most quantum transport effects could be recast into a classical simulation picture of nonlocal collisions"*, cf. [34,35].

2. The Dissipation Function

Consider a system whose "microscopic" state is described by a dynamical system

$$\dot{\mathbf{x}} = \mathbf{V}(\mathbf{x}); \quad \mathbf{x} \in \mathcal{M} \subset \mathbb{R}^n; \quad \mathcal{M} = \text{phase space} \tag{1}$$

Let (1) generate a flow on $\mathcal{M}$, i.e., a map $S^t : \mathcal{M} \to \mathcal{M}$, such that $S^t(\mathbf{x})$, $t \in \mathbb{R}$ is the solution of (1) at time t with initial condition $\mathbf{x}$. Further properties of $\mathcal{M}$ and $\mathbf{V}(\mathbf{x})$ will be specified when needed. Let μ_0 be a probability measure absolutely continuous w.r.t. the Lebesgue measure $d\mathbf{x}$, and denote by $f_0(\mathbf{x})$ its density function, i.e., $d\mu_0(\mathbf{x}) = f_0(\mathbf{x})d\mathbf{x}$. A measure μ is called invariant if $\mu(S^{-t}E) = \mu(E)$ for any measurable set $E \subset \mathcal{M}$ and for all $t \in \mathbb{R}$. If μ_0 is not invariant, it evolves under the dynamics, so that at time t it is expressed by $\mu_t(E) = \mu_0(S^{-t}E)$. Provided the flow is sufficiently smooth, μ_t is absolutely continuous, and has density $f_t(\mathbf{x})$ that satisfies [27]:

$$\frac{\partial f_t}{\partial t}(\mathbf{x}) = f_t(\mathbf{x})\,\Omega^{f_t}(\mathbf{x}), \tag{2}$$

where $\Omega^{f_t}(\mathbf{x})$ is the dissipation function, defined by:

$$\Omega^{f_t}(\mathbf{x}) := -\left[\Lambda(\mathbf{x}) + \mathbf{V}(\mathbf{x}) \cdot \nabla \ln f_t(\mathbf{x})\right]. \tag{3}$$

and

$$\Lambda(\mathbf{x}) := \text{div}\,\mathbf{V}(\mathbf{x}), \tag{4}$$

denotes the phase space variation rate. Its negative, $-\Lambda(\mathbf{x})$, is called phase space contraction rate. In the case the dynamical systems represents an equilibrium system, with equilibrium distribution μ_0, so that $f_t(\mathbf{x}) \equiv f_0(\mathbf{x})$, the dissipation function identically vanishes, $\Omega^{f_0}(\mathbf{x}) \equiv 0$. It is obviously true also the converse: if $\Omega^{f_0}(\mathbf{x}) \equiv 0$ then $f_0(\mathbf{x})$ is an invariant density. In the general case, the solution of (2), is given by [27]:

$$f_t(\mathbf{x}) = \exp\left\{\Omega^{f_0}_{-t,0}(\mathbf{x})\right\} f_0(\mathbf{x}), \tag{5}$$

where, for every phase function $\mathcal{O}$ we use the notation

$$\mathcal{O}_{r,s}(\mathbf{x}) = \int_r^s \mathcal{O}\left(S^\tau \mathbf{x}\right) d\tau \quad \text{hence } \Omega^{f_0}_{-t,0}(\mathbf{x}) = \int_{-t}^0 \Omega^{f_0}(S^\tau \mathbf{x})d\tau. \tag{6}$$

Given any observable $\mathcal{O} : \mathcal{M} \to \mathbb{R}$, we define its ensemble average according to the measure $d\mu_t(\mathbf{x}) = f_t(\mathbf{x})d\mathbf{x}$ as

$$\langle \mathcal{O} \rangle_t = \int_{\mathcal{M}} \mathcal{O}(\mathbf{x})f_t(\mathbf{x})d\mathbf{x}. \tag{7}$$

Under suitable conditions, discussed in Section 6, this average can be computed using the dissipation function as follows [27]:

$$\langle \mathcal{O} \rangle_t = \langle \mathcal{O} \rangle_0 + \int_0^t \left\langle (\mathcal{O} \circ S^\tau) \cdot \Omega^{f_0} \right\rangle_0 d\tau. \tag{8}$$

Equation (8) gives the response of an equilibrium system, characterized by the distribution μ_0, invariant for the dynamics $\dot{\mathbf{x}} = \mathbf{V}_0(\mathbf{x})$, which at time 0 is perturbed so that its dynamics becomes Equation (1). The measure μ_0 is not invariant under the new dynamics, hence it evolves together with the averages representing the observables. At time t the initial measure has turned into μ_t, and the initial average $\langle \mathcal{O} \rangle_0$ into $\langle \mathcal{O} \rangle_t$.

In the long time limit, the response may be given by a non equilibrium invariant distribution. As (8) is exact, the existence of such an asymptotic average is equivalent, as a necessary and sufficient condition called ΩT-mixing, to the convergence of the following limit:

$$\lim_{t \to \infty} \int_0^t \left\langle (\mathcal{O} \circ S^\tau) \cdot \Omega^{f_0} \right\rangle_0 d\tau = L_\mathcal{O} \in \mathbb{R}. \tag{9}$$

for any observable $\mathcal{O}$. A sufficient condition for the existence of $L_\mathcal{O}$ is [26]:

$$\left\langle (\mathcal{O} \circ S^t) \, \Omega^{f_0} \right\rangle_0 = o(t^{-1}), \tag{10}$$

A further useful asymptotic property is obtained by considering two observables $\mathcal{O}$ and $\mathcal{Q}$. In this case, we say that the system is *T*-mixing [26] if

$$\lim_{t \to \infty} [\langle (\mathcal{O} \circ S^t) \mathcal{Q} \rangle_0 - \langle (\mathcal{O} \circ S^t) \rangle_0 \langle \mathcal{Q} \rangle_0] = 0. \tag{11}$$

We observe that setting $\mathcal{Q} = \Omega^{f_0}$ in (11) we obtain a weak form of ΩT-mixing, namely

$$\lim_{t \to \infty} \left\langle (\mathcal{O} \circ S^t) \, \Omega^{f_0} \right\rangle_0 = 0, \tag{12}$$

given that one has, cf. Section 6 and [27]:

$$\left\langle \Omega^{f_0} \right\rangle_0 = 0. \tag{13}$$

Obviously, *T*-mixing implies weak ΩT-mixing. We observe that these properties may hold on a specific set of observables for a given f_0 but not for other initial distributions.

A fundamental difference between *T*-mixing and the mixing property of standard ergodic theory is that *T*-mixing concerns a the decay of correlations w.r.t. the initial state μ_0, hence it may properly describe the irreversible relaxation of a given system to a stationary state. Mixing, instead, corresponds to a decay of correlations within a given steady state, since it refers to an invariant measure [5]. We refer to Refs. [5,27] for a broader discussion of these issues.

2.1. Examples

The physical meaning of the Dissipation Function can be appreciated by considering some special choices of vector field $\mathbf{V}$, and of initial probability density f_0. First, consider the NEMD systems for which Ω has been introduced, to see that in such cases it is directly related to the instantaneous dissipative flux. The simplest of these systems have the following equations of motion:

$$\begin{aligned} \dot{\mathbf{q}}_i &= \mathbf{p}_i/m + \mathbf{C}_i(\Gamma) \cdot \mathbf{F}_e, \quad i = 1, ..., N \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i + \mathbf{D}_i(\Gamma) \cdot \mathbf{F}_e - \alpha \mathbf{p}_i \end{aligned} \tag{14}$$

where $\Gamma = (\mathbf{q}, \mathbf{p})$ collectively represents all coordinates $\mathbf{q}$ and momenta $\mathbf{p}$ of the N particles, and $\mathbf{F}_e$ is the external driving field, coupled to the particles via the "charges" $\mathbf{C}_i(\Gamma)$ and $\mathbf{D}_i(\Gamma)$. The term α is a Lagrange multiplier meant to implement some constraint, so that the system reaches a steady state despite the driving field. In the NEMD literature α is said to implement a deterministic, time reversible thermostat [6]. If this thermostat is set to 0, one obtains the adiabatic equations of motion:

$$\begin{aligned}\dot{\mathbf{q}}_i &= \mathbf{p}_i/m + \mathbf{C}_i(\Gamma) \cdot \mathbf{F}_e\,, \quad i = 1, ..., N \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i + \mathbf{D}_i(\Gamma) \cdot \mathbf{F}_e \end{aligned} \tag{15}$$

which imply a continuous growth of energy. Then, one must assume that the support of f_0 is densely explored by a single equilibrium trajectory [30].

Denoting by Φ the internal interaction potential energy of the N particles, the dissipative flux, $\mathbf{J}$, is obtained from the time derivative of the internal energy

$$H_0 = \sum_{i=1}^{N} \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{2m} + \Phi(\mathbf{q})$$

under the dynamics (15), which represents how much energy has to be dissipated in order to keep the steady state. The result is:

$$\dot{H}_0^{ad} = \sum_{i=1}^{N} \left(\frac{\mathbf{p}_i}{m} \cdot \mathbf{D}_i - \mathbf{F}_i \cdot \mathbf{C}_i \right) \cdot \mathbf{F}_e \equiv -\mathbf{J}V \cdot \mathbf{F}_e \tag{16}$$

where V is the volume of the system.

Consider a few common examples.

2.1.1. Microcanonical Ensemble

For the isoenergetic dynamics, in which the internal energy is conserved, i.e., H_0 is fixed and $\dot{H}_0 = 0$, the equilibrium ensemble f_0 is microcanonical. Therefore, for an equilibrium Hamiltonian N particle system in d dimensions, one has [30]:

$$\alpha(\Gamma) = -\frac{\mathbf{J}(\Gamma)V \cdot \mathbf{F}_e}{2K(\Gamma)} \tag{17}$$

where K is the kinetic energy of the N particles, while

$$\Lambda(\Gamma) = \sum_{i=1}^{N} \left(\frac{\partial \mathbf{C}_i(\Gamma)}{\partial \mathbf{q}_i} + \frac{\partial \mathbf{D}_i(\Gamma)}{\partial \mathbf{p}_i} \right) \cdot \mathbf{F}_e - dN\alpha(\Gamma) + O\left(\frac{1}{N}\right) = -dN\alpha(\Gamma) + O\left(\frac{1}{N}\right) \tag{18}$$

as in the original paper [36]. Then, one may write:

$$\Omega_{t,t+\tau}^{(0)}(\Gamma) = \ln \frac{f_0(S^t\Gamma)}{f_0(S^{t+\tau}\Gamma)} - \Lambda_{t,t+\tau}(\Gamma) = -\Lambda_{t,t+\tau}(\Gamma)$$

As even the $O(1/N)$ correction in Equation (18) is proportional to α, one observes that in this case both Ω and Λ are proportional to the rate of energy dissipation (17).

2.1.2. Gaussian Isokinetic System

In this case the kinetic energy K is fixed, and one may formally write $K = (dN - d - 1)k_B T = (dN - d - 1)/\beta$, if also the center of mass is fixed, to prevent contributions from drift, so that:

$$\alpha(\Gamma) = -\frac{(\dot{H}_0(\Gamma) + \mathbf{J}(\Gamma)V \cdot \mathbf{F}_e)\beta}{dN - d - 1} \tag{19}$$

In turn, Λ is still expressed by (18), with a different $O(1/N)$ term. In this case, f_0 writes:

$$f_0(\Gamma) = \frac{1}{Q} e^{-\beta H_0} \delta(K(\Gamma) - K_0)$$

where Q is the normalization constant. The result for the dissipation function is:

$$\overline{\Omega}_{0,\tau}(\Gamma) = \frac{\beta}{\tau} \int_0^\tau \dot{H}_0(S^s\Gamma)ds - \overline{\Lambda}_{0,\tau}(\Gamma) = -\overline{(\mathbf{J}\cdot\mathbf{F}_e)}_{0,\tau}V\beta + O\left(\frac{1}{N}\right) \tag{20}$$

where the bar over the observables denotes time average. Combining this with Equations (18) and (19), one obtains

$$\sigma(\Gamma) = -\mathbf{J}(\Gamma)\cdot\mathbf{F}_eV\beta + O\left(\frac{1}{N}\right) . \tag{21}$$

where σ is the entropy production, if local thermodynamic equilibrium is verified. We observe that $-\Lambda$ and Ω differ by a term proportional to the fluctuating $\dot{H}$. If these fluctuations average to zero, the averages of $-\Lambda$ and Ω coincide, however, depending on how the respective fluctuations are distributed, various relations may apply to one and not to the other. For instance, the steady state fluctuation relation may hold for Ω but not for Λ, [37,38].

2.1.3. Nosé-Hoover Thermostat

In this case, there is one extra equation to be added to the equations of motion (14):

$$\dot{\zeta} = \frac{1}{\tau}\left(2K(\Gamma) - dNk_BT\right) \tag{22}$$

where τ sets the relaxation time scale of the thermostat, and T is the imposed average temperature. Then, we may write:

$$\zeta(\Gamma) = -\left(\dot{H}_0(\Gamma) + \frac{\mathbf{J}(\Gamma)V\cdot\mathbf{F}_e}{2K(\Gamma)}\right) \tag{23}$$

and for unperturbed Hamiltonian dynamics, one may write:

$$\Lambda(\Gamma) = \nabla\cdot\dot{\Gamma} + \frac{\partial\dot{\zeta}}{\partial\zeta} = -dN\zeta . \tag{24}$$

In this case, f_0 is given by

$$f_0(\Gamma) = \frac{1}{M} e^{-\beta(H_0 + \frac{1}{2}Q\zeta^2)}$$

where M is the normalization. It follows that:

$$\frac{f_0(\Gamma)}{f_0(S^\tau\Gamma)} = e^{\beta\int_0^\tau(\dot{H}_0(S^s\Gamma) + Q\zeta(S^s\Gamma)\dot{\zeta}(S^s\Gamma))ds}$$

which implies:

$$\Omega(\Gamma) = -\mathbf{J}(\Gamma)\cdot\mathbf{F}_eV\beta . \tag{25}$$

We conclude that for equilibrium Hamiltonian dynamics, the driven nonequilibrium evolution is characterized by the dissipation function, which equals the entropy production, in case local thermodynamic equilibrium holds. Additionally, in this case $-\Lambda$ and Ω differ by a term proportional to $\dot{H}_0$, and the same remarks concluding the previous subsection equally apply.

2.1.4. Canonical Ensemble for Generic Dynamics

The canonical ensemble describes the equilibrium state of a Hamiltonian particle systems in contact with a heat bath at temperature $k_BT = \beta^{-1}$, with which it can exchange energy but not

matter. Here, we consider a general case in which the vector field $\mathbf{V}(\mathbf{x})$ is not necessarily Hamiltonian, but there still is an energy function H. In addition, we assume that the initial distribution has the same form of the canonical Gibbs density:

$$f_0(\mathbf{x}) = \frac{e^{-\beta H(\mathbf{x})}}{Z_\beta},$$

(26)

Z_β being the normalization constant. Then, the Dissipation Function reads:

$$\Omega^{f_0}(\mathbf{x}) = \beta \dot{H}(\mathbf{x}) - \Lambda(\mathbf{x}),$$

(27)

where we use the notation $\dot{H}(\mathbf{x}) = \mathbf{V}(\mathbf{x}) \cdot \nabla H(\mathbf{x})$. The Dissipation Function Ω^{f_0} is not necessarily zero, in this case, and the density evolves as follows, cf. (5):

$$f_t(\mathbf{x}) = f_0(\mathbf{x}) \exp\left\{\beta(H(\mathbf{x}) - H(S^{-t}\mathbf{x})) - \Lambda_{-t,0}(\mathbf{x})\right\} = \frac{\exp\left\{-\beta H(S^{-t}\mathbf{x})\right\}}{Z_t(\mathbf{x})}.$$

where: $Z_t^{-1}(\mathbf{x}) = \exp\{-\Lambda_{-t,0}(\mathbf{x})\}Z_\beta^{-1}$. In the case of Hamiltonian dynamics, $\mathbf{x} = (\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2n}$, $\mathbf{V}(\mathbf{x}) = \mathbf{J}\nabla H(\mathbf{x})$ where $\mathbf{J}$ is a symplectic matrix, we have $\dot{H}(\mathbf{x}) = \mathbf{V}(\mathbf{x}) \cdot \nabla H(\mathbf{x}) \equiv 0$ and $\Lambda(\mathbf{x}) \equiv 0$. Thus, in the equilibrium canonical ensemble, where (26) is invariant and the phase space volume is conserved, the dissipation function is identically zero $\Omega^{f_0}(\mathbf{x}) \equiv 0$.

The extended canonical equilibrium distribution of the Nosé–Hoover thermostatted system can be framed in this way: its dynamics includes a degree of freedom ζ that represents the heat bath, and system plus bath constitute a canonical equilibrium state. For instance, consider a system made of N identical noninteracting one-dimensional oscillators in contact with a bath at temperature T, implemented by a single Nosé–Hoover thermostat. This thermostat introduces an indirect coupling among the oscillators, which however does not represent a proper interaction leading to LTE. We illustrate these facts considering both linear and nonlinear oscillators, i.e., potential energies $U(q) = -\omega_0^2 q^{\alpha+1}/(\alpha + 1)$, with $\alpha = 1$ and $\alpha = 3$. The equations of motion are

$$\begin{cases} \dot{q}_i = \dfrac{p_i}{m}, \\ \dot{p}_i = -m\,\omega_0^2\,q_i^\alpha - \zeta p_i, \quad i = 1, \dots, N, \\ \dot{\zeta} = \dfrac{1}{\tau}\left(\dfrac{2K(p)}{T} - 1\right), \end{cases}$$

(28)

where T is the target kinetic energy per particle,

$$K(p) = \frac{1}{2mN}\sum_{i=1}^{N} p_i^2$$

is the kinetic energy per particle, and τ the relaxation time. In order to elucidate the non-thermodynamic nature of the model (28), we simulated a system of $N = 30$ oscillators in the linear case, $\alpha = 1$, with $m = 1$, $\omega_0^2 = 2$, $\tau = 2$ and $T = 100$. As expected because of the improper particles coupling expressed by ζ, we observe no equipartition; cf. Figure 1, which shows that the time average $\overline{K}_i$ of the kinetic energies of single particles substantially change with their label i, while the time average $\overline{K}$ of the total is equal to T within numerical errors. The same happens also with $\alpha = 3$.

To promote LTE in this system, we introduce nearest neighbour interactions as:

$$\dot{p}_i = -m\,\omega_0^2\,q_i^\alpha + k(q_{i-1} - 2q_i + q_{i+1}) - \zeta p_i, \quad i = 1, \dots, N,$$

(29)

in place of the second of (28), taking $k > 0$ as the coupling strength, and $q_0 = q_n$, $q_{n+1} = q_0$ as boundary conditions. Unlike the previous non-interacting model, the linear ($\alpha = 1$) and nonlinear ($\alpha = 3$) cases behave in a different way. While still absent in the linear case, equipartition is established

with the cubic local field. Figure 2 shows the time averages $\frac{1}{J}\sum_{j=1}^{J} p_i^2(t_j)$ of the kinetic energies of some particles sampled form a long trajectory of the coupled system with $N = 10$. We consider the cases with $k = 1, 2, 5$ in order to make evident the role of the coupling. These simulations show that increasing the particles coupling k speeds up the equipartition process. Changing the oscillators frequency, instead, has no effect on equipartition, cf. Figure 1.

Now, focusing on the energy $H_0 = \frac{1}{2m}\sum_{i=1}^{N} p_i^2 + \frac{m}{2}\omega_0^2 \sum_{i=1}^{N} q_i^2$ of the system of oscillators alone, i.e., without the bath, and taking f_0 as

$$f_0(\Gamma) = \frac{1}{Z_0} e^{-\beta H_0(\Gamma)} \tag{30}$$

in order to define a kind of reduced dissipation function for the dynamics with the bath, one obtains:

$$\widehat{\Omega}_0^{f_0} = \zeta\left(N - \frac{\beta}{m}\sum_{i=1}^{N} p_i^2\right) = N\zeta\left(1 - 2T^{-1}K(p)\right) = -\Theta^2 N\zeta\dot{\zeta} = -N\Theta^2 \frac{d}{dt}\left(\frac{\zeta^2}{2}\right),$$

which is not identically zero even in the equilibrium state. There is no immediate physical interpretation of the quantity $\widehat{\Omega}_0^{f_0}$, but it clearly shows that, in this framework, the thermostat has to be included to properly speak of equilibrium.
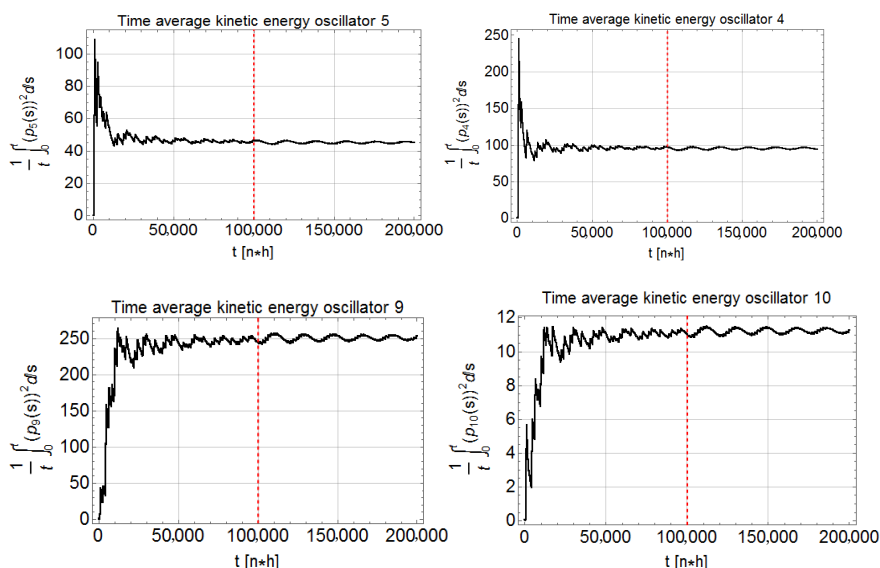

Figure 1. System of linear oscillators, $\alpha = 1$, subjected to a Nosé–Hoover thermostat, (28), with $N = 30$, $m = 1$, $\omega_0^2 = 2$, $\tau = 2$ and $T = 100$. Time average of kinetic energy of oscillators $i = 4, 5, 9, 10$ versus time. The dependence of the asymptotic values on i demonstrate the violation of equipartition. The dashed vertical line marks the instant of time at which the constant ω_0^2 is changed to $\omega_0^2 + \delta\omega_0^2$ to check whether the oscillator "temperature" also changes. It appears that it does not.
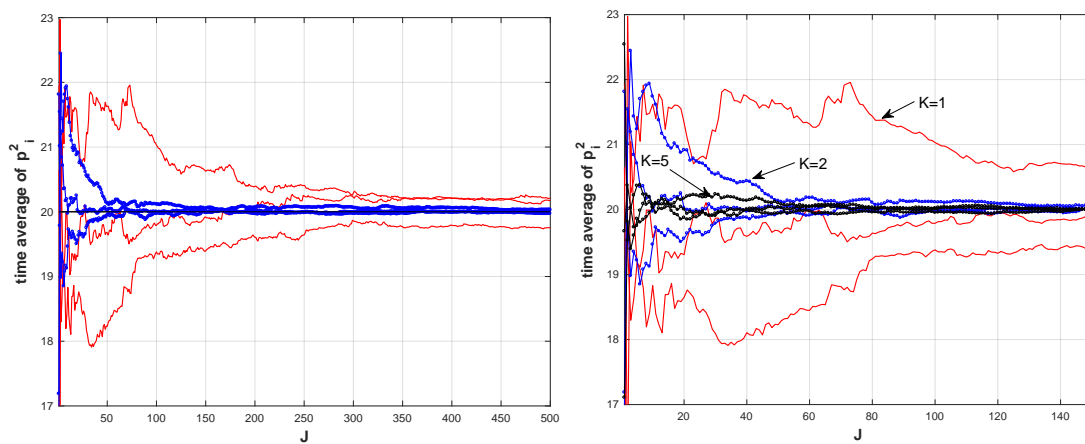
Figure 2. System with $N = 10$ coupled oscillators, with $\alpha = 3$, $m = 1$, $\omega_0^2 = 1$. The temperature forced by the Nosé–Hoover thermostat is $T = 20$, and the response time $\tau = 1$. The plots represent the time averages $\frac{1}{J} \sum_{j=1}^{J} p_i^2(t_j)$, of the kinetic energies of three particles ($i = 1, 4, 8$) sampled at $t_j = 10^3 \cdot j$, for increasing J. In the left panel, where $J = 500$, the outer continuous (red online) lines correspond to the coupling parameter $k = 1$, while the inner dotted lines (blue online) correspond to $k = 2$. In the right panel the focus is on a shorter interval, $J = 150$, where the relaxation phase can be better appreciated. In this case the outer lines (red online) correspond to $k = 1$, the middle lines (blue online) correspond to $k = 2$, the inner line (black) correspond to $k = 5$. Larger coupling enhances the convergence rate towards equipartition, as expected.

2.1.5. Hamiltonian and Gradient Vector Field with Canonical and Gaussian f_0

In the case of Hamiltonian dynamics perturbed by a gradient term [39]:

$$\mathbf{V}(\mathbf{x}) = \mathbf{J}\nabla H(\mathbf{x}) - \nabla L(\mathbf{x}), \tag{31}$$

where the unperturbed equilibrium density should be chosen consistently with the constraints that remain in place when the perturbation is switched on, in order to implement the above-mentioned ergodic consistency condition. For instance, if the system temperature remains fixed, one may take the canonical distribution, $f_0(\mathbf{x}) = e^{-\beta H(\mathbf{x})}/Z$, that implies:

$$\Omega^{f_0}(\mathbf{x}) = \Delta L(\mathbf{x}) - \beta \nabla L(\mathbf{x}) \cdot \nabla H(\mathbf{x}),$$

$\Delta L = -\Lambda$ being the Laplacian of L. One observes that Ω^{f_0} only depends on L if ∇L is orthogonal to ∇H. If ∇L is in addition a Hamiltonian perturbation, the perturbed steady state remains equal to the unperturbed equilibrium. Now, taking an initial uncorrelated Gaussian distribution:

$$f_0(\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x_i - \mu_i)^2}{2\sigma_i^2}} \tag{32}$$

one obtains:

$$\Omega^{f_0} = -\Lambda(\mathbf{x}) + \sum_{i=1}^{n} V_i(\mathbf{x}) \frac{x_i - \mu_i}{\sigma_i^2} = \Delta L(x) + \sum_{i=1}^{n} \left(\sum_{j=i}^{N} J_{ij} \frac{\partial H(x)}{\partial x_j} - \frac{\partial L(x)}{\partial x_i} \right) \frac{x_i - \mu_i}{\sigma_i^2}, \tag{33}$$

J_{ij} being the entries of $\mathbf{J}$. Again, a Hamiltonian perturbation such that

$$\sum_{j=i}^{N} J_{ij} \frac{\partial H(x)}{\partial x_j} - \frac{\partial L(x)}{\partial x_i} = 0, \quad \forall i = 1, ..., N \tag{34}$$

implies that the Gaussian (32) is invariant under the dynamics. These two examples constitute elementary warnings against too quick thermodynamic interpretations of generic particle systems, much as it may be interesting in certain circumstances [40].

2.2. The Dissipation Function and the Phase Space Contraction Rate

The Dissipation Function satisfies the Non Equilibrium Partition identity [27]:

$$\left\langle \exp\left(-\Omega_{-t,0}^{f_0}\right)\right\rangle_0 = 1 \tag{35}$$

Using Jensen inequality, we can also write:

$$\exp\left[-\left\langle \Omega_{-t,0}^{f_0}\right\rangle_0\right] \leq 1 \tag{36}$$

obtaining:

$$\left\langle \Omega_{-t,0}^{f_0}\right\rangle_0 \geq 0$$

in accordance with both our physical intuition on the dissipation of energy in macroscopic systems and with the interpretation of this quantity as a Kullback–Leibler divergence between the initial and the evolved probability distribution in phase space [27].

One motivation to derive relations like these, see e.g., Refs. [27,30] where various others are derived, resides in the fact that they are only based on the dynamics and on the initial probability distribution, whatever phenomenon they represent, hence they hold very generally. At the same time, this also means that blind thermodynamic interpretations are to be avoided, although no difficulty arises from a purely dynamical systems standpoint.

The distinction between the dissipation function and the phase space contraction rate $-\Lambda$ becomes particularly evident when considering particle systems. In fact, the steady state phase space average of $-\Lambda$ equals the steady state entropy production of certain NEMD models [6,19], and it can be expressed as:

$$e_p = -\int_{\mathcal{M}} \Lambda(\mathbf{x}) d\mu_{(}\mathbf{x}), \tag{37}$$

μ being the corresponding invariant measure, and e_p the steady state entropy production or the energy dissipation. On these grounds, Gallavotti and Cohen proposed $-\Lambda$ as the steady state definition of entropy production tout court, [41], followed by Ruelle [23]. Earlier, a similar idea had been contemplated by Andrey [20], who referred to the variation in time of the Gibbs entropy. The Gibbs entropy, however, is not directly related, in general, to the entropy of a nonequilbrium system [42–44]. (Note: this statement does not concern the formally analogous definition of entropy common in the theory of stochastic processes, since they are not defined in phase space but in the space of observables.) Therefore, $-\Lambda$ cannot be directly related, in general, to the entropy production, when considering time dependent states or steady state fluctuations, cf. Section 2.1 and Refs. [27,30,38].

On the other hand, it is interesting to investigate both $-\Lambda$ and Ω as properties of dynamical systems, in a kind of extension of the thermodynamic formalism [33], that refers the non-equilibrium language.

Given an initial absolutely continuous probability measure $d\mu_0(\mathbf{x}) = f_0(\mathbf{x})d\mathbf{x}$ on the phase space $\mathcal{M}$, and letting μ_t be the evolution of μ_0 at time t, the corresponding average of the dissipation function is:

$$\langle \Omega^{f_0}\rangle_t = \int_{\mathcal{M}} \Omega^{f_0} d\mu_t = -\langle \Lambda\rangle_t - \langle V\cdot\nabla\ln f_0\rangle_t \tag{38}$$

which in the case of a uniform density f_0 yields:

$$\langle \Omega^{f_0}\rangle_t = -\langle \Lambda\rangle_t \tag{39}$$

In this case, even Ω^{f_0} does not need to represent any entropy production rate: the equilibrium dynamics would have at least to be phase space volumes preserving, for Ω^{f_0} to be related to the dissipation of energy [30]. Using Equation (8) and the fact that $\langle \Omega^{f_0} \rangle_0 = 0$, we may also write:

$$- \langle \Lambda \rangle_t = - \int_0^t \left\langle \Omega^{f_0} \left(\Lambda \circ S^\tau \right) \right\rangle_0 d\tau \tag{40}$$

In the case of a generic probability density f_0, the exact response formula (8) and (40) yield:

$$\left\langle \left(-\Lambda - \Omega^{f_0} \right) \circ S^t \right\rangle_0 = -\langle \Lambda \rangle_t - \left\langle \Omega^{f_0} \right\rangle_t = \langle V \cdot \nabla \ln f_0 \rangle_0 + \int_0^t \left\langle \Omega^{f_0} \left[(V \cdot \nabla \ln f_0) \circ S^\tau \right] \right\rangle_0 d\tau \tag{41}$$

which quantifies the difference between the mean phase space contraction rate and the energy dissipation in time. In the case ΩT-mixing, Equation (9), holds, t can be replaced by ∞ in (41), obtaining the asymptotic difference between the two quantities:

$$\lim_{t \to \infty} \left[\langle -\Lambda \rangle_t - \left\langle \Omega^{f_0} \right\rangle_t \right] = \langle V \cdot \nabla \ln f_0 \rangle_0 + \int_0^\infty \left\langle \Omega^{f_0} \left[(V \cdot \nabla \ln f_0) \circ S^\tau \right] \right\rangle_0 d\tau. \tag{42}$$

From the definition of dissipation function Ω^{f_0}, i.e., Equation (3) with $t = 0$, one also observes that

$$\left(-\Lambda - \Omega^{f_0} \right) \circ S^s = \frac{d}{dt} \ln f_0 \circ S^t \Big|_{t=s} \tag{43}$$

Therefore, the time integrated difference of $-\Lambda$ and Ω^{f_0} merely amounts to a boundary term:

$$\int_0^t \left(-\Lambda - \Omega^{f_0} \right) (S^s \mathbf{x}) ds = \ln f_0(\mathbf{x}) - \ln f_0(S^t \mathbf{x}) \tag{44}$$

which, however, may bear noticeably affect the statistic of fluctuations: the distribution of such boundary terms may, for instance, imply the validity of the FR for Ω^{f_0} and not for $-\Lambda$, cf. [37,38,45].

3. Exact Response and GK Formalism

In this section, we compare the classical linear response theory, or Green–Kubo formalism, with the exact response represented by Equation (8). Consider a perturbed vector field

$$\mathbf{G}_\epsilon(\mathbf{x}) = \mathbf{G}_0(\mathbf{x}) + \epsilon\, \mathbf{G}_p(\mathbf{x}), \tag{45}$$

where ϵ may or may not be small. Then, the Dissipation Function can be written as:

$$\Omega_{\mathbf{G}_\epsilon}^{f_0}(\mathbf{x}) = \Omega_{\mathbf{G}_0}^{f_0}(\mathbf{x}) + \epsilon\, \Omega_{\mathbf{G}_p}^{f_0}(\mathbf{x}). \tag{46}$$

and the evolution of a probability density $f_0(\mathbf{x})$ is given by

$$f_t(\mathbf{x}; \epsilon) = f_0(\mathbf{x}) \exp \left[\int_{-t}^0 d\tau\, \Omega_{\mathbf{G}_0}^{f_0}(S_\epsilon^\tau \mathbf{x}) \right] \exp \left[\epsilon \int_{-t}^0 d\tau\, \Omega_{\mathbf{G}_p}^{f_0}(S_\epsilon^\tau \mathbf{x}) \right]. \tag{47}$$

where S_ϵ^t denotes the flow of the vector field $\mathbf{G}_\epsilon(\mathbf{x})$. Thus, the exact response equation (8) reads

$$\langle \mathcal{O} \rangle_{t;\epsilon} := \int \mathcal{O}(\mathbf{x}) f_t(\mathbf{x}; \epsilon) d\mathbf{x}$$
$$= \langle \mathcal{O} \rangle_0 + \int_0^t \left\langle (\mathcal{O} \circ S_\epsilon^\tau \mathbf{x}) \cdot \Omega_{\mathbf{G}_0}^{f_0}(\mathbf{x}) \right\rangle_0 d\tau + \epsilon \int_0^t \left\langle (\mathcal{O} \circ S_\epsilon^\tau \mathbf{x}) \cdot \Omega_{\mathbf{G}_p}^{f_0}(\mathbf{x}) \right\rangle_0 d\tau. \tag{48}$$

In case $\mathbf{G}_0$ derives from a Hamiltonian H_0, the unperturbed dynamics preserves phase space volumes, and one can write:

$$\Omega_{\mathbf{G}_0}^{f_0} = -\mathbf{G}_0 \cdot \nabla \ln f_0 = -\frac{1}{f_0}[f_0, H_0] = \frac{1}{f_0}\mathcal{L}_0 f_0 \tag{49}$$

where $\mathcal{L}_0 = [H_0, \cdot]$ is the Liouvillean corresponding to H_0 and $[\cdot, \cdot]$ denotes the associated Poisson brackets. Then, we may define the perturbation operator $\mathcal{L}_p$ as:

$$\Omega_{\mathbf{G}_p}^{f_0} = -\mathbf{G}_p \cdot \nabla \ln f_0 = -\frac{1}{f_0}[f_0, H_p] = \frac{1}{f_0}\mathcal{L}_p f_0 \tag{50}$$

so that the (total) dissipation function can be written as:

$$\Omega_{\mathbf{G}_\epsilon}^{f_0} = \frac{1}{f_0}(\mathcal{L}_0 + \epsilon \mathcal{L}_p)f_0, \tag{51}$$

Substituting (51) in (48) yields:

$$
\begin{aligned}
\langle \mathcal{O} \rangle_{t;\epsilon} &= \langle \mathcal{O} \rangle_0 + \int_0^t \left\langle (\mathcal{O}(S_\epsilon^\tau \mathbf{x}) \frac{1}{f_0(\mathbf{x})} \mathcal{L}_0 f_0(\mathbf{x}) \right\rangle_0 d\tau + \epsilon \int_0^t \left\langle (\mathcal{O}(S_\epsilon^\tau \mathbf{x}) \frac{1}{f_0(\mathbf{x})} \mathcal{L}_p f_0(\mathbf{x}) \right\rangle_0 d\tau \\
&= \langle \mathcal{O} \rangle_0 + \int_0^t d\tau \int_{\mathcal{M}} d\mathbf{x}\, \mathcal{O}(S_\epsilon^\tau \mathbf{x}) \mathcal{L}_0 f_0(\mathbf{x}) + \epsilon \int_0^t d\tau \int_{\mathcal{M}} d\mathbf{x}\, \mathcal{O}(S_\epsilon^\tau \mathbf{x}) \mathcal{L}_p f_0(\mathbf{x})
\end{aligned}
\tag{52}
$$

and in case the initial distribution $f_0(\mathbf{x})$ is the equilibrium state of the unperturbed system, so that $\mathcal{L}_0 f_0 \equiv 0$, one obtains:

$$\langle \mathcal{O} \rangle_{t;\epsilon} = \langle \mathcal{O} \rangle_0 + \epsilon \int_0^t d\tau \int_{\mathcal{M}} d\mathbf{x}\, \mathcal{O}(S_\epsilon^\tau \mathbf{x})\, \mathcal{L}_p f_0(\mathbf{x}). \tag{53}$$

Provided the response Equation (8) applies, the formulae (52) and (53) for the evolution of phase space averages are both exact: there is no assumption on the size of ϵ and no approximation has been made. An approximate formula can be obtained in the small-ϵ limit, starting from Equation (47), by truncating to first order the exponential expansion, and recalling that $\Omega_{\mathbf{G}_0}^{f_0}(\mathbf{x}) \equiv 0$. This yields:

$$\tilde{f}_t(\mathbf{x};\epsilon) := f_0(\mathbf{x}) + \epsilon f_0(\mathbf{x}) \int_{-t}^0 d\tau\, \Omega_{\mathbf{G}_p}^{f_0}(S_\epsilon^\tau \mathbf{x}) = f_0(\mathbf{x}) + \epsilon f_0(\mathbf{x}) \int_0^t d\tau\, \Omega_{\mathbf{G}_p}^{f_0}(S_\epsilon^{-\tau} \mathbf{x}). \tag{54}$$

where the tilde symbol stresses that this is but an approximation of the probability density at time t, which is not guaranteed to be a probability density itself. Nevertheless, multiplying (54) by $\mathcal{O}$ and integrating over the phase space, one gets:

$$\langle \mathcal{O} \rangle_{\widetilde{t;\epsilon}} := \int_{\mathcal{M}} \mathcal{O}(\mathbf{x})\tilde{f}_t(\mathbf{x};\epsilon)d\mathbf{x} = \langle \mathcal{O} \rangle_0 + \epsilon \int_0^t d\tau \left\langle \mathcal{O}(\mathbf{x})\Omega_{\mathbf{G}_p}^{f_0}(S_\epsilon^{-\tau} \mathbf{x}) \right\rangle_0 \tag{55}$$

In order to keep the truncation error bounded in time, one may assume that the time integral exists. For instance, $\left\langle \mathcal{O}(\mathbf{x})\Omega_{\mathbf{G}_p}^{f_0}(S_\epsilon^{-t} \mathbf{x}) \right\rangle_0$ maybe of order $o(t^{-1})$, which is reminiscent of the ΩT-mixing condition. The difference between the exact response formula, (48), and the approximate one, (55), amounts to:

$$\langle \mathcal{O} \rangle_{t;\epsilon} - \langle \mathcal{O} \rangle_{\widetilde{t;\epsilon}} = \epsilon \int_0^t d\tau \left\{ \left\langle \mathcal{O}(S_\epsilon^\tau \mathbf{x}) \cdot \Omega_{\mathbf{G}_p}^{f_0}(\mathbf{x}) \right\rangle_0 - \left\langle \mathcal{O}(\mathbf{x})\Omega_{\mathbf{G}_p}^{f_0}(S_\epsilon^{-\tau} \mathbf{x}) \right\rangle_0 \right\}, \tag{56}$$

that, in general, does not vanish, and may grow large in time.

May the difference of responses Equation (56) become negligible, in some limit? Applying the transformation $S_\epsilon^{-\tau}\mathbf{x} = \mathbf{y}$ to the second integrand of (56), one obtains:

$$\left\langle \mathcal{O}(\mathbf{x})\Omega_{\mathbf{G}_p}^{f_0}(S_\epsilon^{-\tau}\mathbf{x}) \right\rangle_0 = \int_{\mathcal{M}} \mathcal{O}(\mathbf{x})\Omega_{\mathbf{G}_p}^{f_0}(S_\epsilon^{-\tau}\mathbf{x})f_0(\mathbf{x})d\mathbf{x} = \int_{\mathcal{M}} \mathcal{O}(S_\epsilon^{\tau}\mathbf{y})\Omega_{\mathbf{G}_p}^{f_0}(\mathbf{y})f_0(S_\epsilon^{\tau}\mathbf{y})\left|J_{S_\epsilon^{\tau}}(\mathbf{y})\right|d\mathbf{y}, \quad (57)$$

where $\left|J_{S_\epsilon^{\tau}}(\mathbf{y})\right| = 1$ since we are dealing with Hamiltonian dynamics–being $\left|J_{S_\epsilon^{\tau}}(\mathbf{y})\right|$ the Jacobian determinant. Thus, Equation (56) can be rewritten as:

$$\langle \mathcal{O} \rangle_{t;\epsilon} - \langle \mathcal{O} \rangle_{\widetilde{t;\epsilon}} = \epsilon \int_0^t d\tau \int_{\mathcal{M}} \mathcal{O}(S_\epsilon^{\tau}\mathbf{x})\Omega_{\mathbf{G}_p}^{f_0}(\mathbf{x})\left[f_0(\mathbf{x}) - f_0(S_\epsilon^{\tau}\mathbf{x})\right]d\mathbf{x} \quad (58)$$

For every finite t, this quantity vanishes in the $\epsilon \to 0$ limit but, unlike the case of the Green–Kubo theory of linear response, the flow used in (58) is the perturbed one, S_ϵ^{τ}, not the unperturbed S_0^{τ}.

As a matter of fact, the success of the Green–Kubo theory is based on its ability to compute the response by averaging with respect to the equilibrium ensemble for the unperturbed dynamics, and by simulating such an unperturbed dynamics [46]. The procedure, illustrated in many textbooks, including [46,47], uses the canonical ensemble for the unperturbed Hamiltonian H_0, $Z_0^{-1}\exp\{-\beta H_0\}$, and assumes that the parameter ϵ is the magnitude of and external force $\mathbf{F}_e$. Within the framework of NEMD and Equation (14), one starts from the unthermostatted, $\alpha = 0$ situation. Without loss of generality, the external driving $\mathbf{F}_e$ can be written as a scalar quantity, since its direction can be embedded in the coefficients $\mathbf{C}_i$ and $\mathbf{D}_i$. Then, the unthermostatted perturbed equations are derived from a Hamiltonian, $H = H_0 + H_p$. Writing $\mathbf{x} = (\mathbf{q}, \mathbf{p})$ to distinguish configuration from momenta, $\mathbf{G}_p$ is obtained deriving H_p:

$$C_i F_e = \frac{\partial H_p}{\partial p_i}; \quad D_i F_e = -\frac{\partial H_p}{\partial q_i} \quad (59)$$

and the following holds:

$$0 = \Lambda = \frac{\partial \dot{\mathbf{q}}}{\partial \mathbf{q}} + \frac{\partial \dot{\mathbf{p}}}{\partial \mathbf{p}} = \sum_{i=1}^{N}\left\{\frac{\partial C_i}{\partial q_i} + \frac{\partial D_i}{\partial p_i}\right\}F_e, \quad \text{hence} \quad \sum_{i=1}^{N}\left\{\frac{\partial C_i}{\partial q_i} + \frac{\partial D_i}{\partial p_i}\right\} = 0. \quad (60)$$

In this case, one has:

$$\Omega_{\mathbf{G}_p}^{f_0} = -\mathbf{G}_p \cdot \nabla \ln f_0 = \frac{1}{f_0}\frac{\partial f_0}{\partial H_0}\mathbf{G}_p \cdot \nabla H_0 = -\beta \sum_{i=1}^{N}\left[\mathbf{C}_i\frac{\partial \dot{H}_0}{\partial \mathbf{q}_i} + \mathbf{D}_i\frac{\partial \dot{H}_0}{\partial \mathbf{p}_i}\right]F_e = -\beta V J F_e \quad (61)$$

which equals Equation (16), and shows again that $\Omega_{\mathbf{G}_\epsilon}^{f_0}$, not $-\Lambda$, represents that energy dissipation rate. Now, following the Green–Kubo approach [46,47], which rests on the fact that

$$\left[H_0, H_p\right] = \left[H, H_p\right], \quad (62)$$

one may use the unperturbed dynamics, and eventually write:

$$\langle \mathcal{O} \rangle_{\widetilde{t;F_e}} = \langle \mathcal{O} \rangle_0 - \beta F_e \int_0^t d\tau \left\langle \left(\mathcal{O} \circ S_0^{t-\tau}\right)\sum_{i=1}^{N}\left[\mathbf{C}_i\frac{\partial \dot{H}_0}{\partial \mathbf{q}_i} + \mathbf{D}_i\frac{\partial \dot{H}_0}{\partial \mathbf{p}_i}\right]\right\rangle_0 \quad (63)$$

$$= \langle \mathcal{O} \rangle_0 - \beta V F_e \int_0^t d\tau \left\langle J\left(\mathcal{O} \circ S_0^{t-\tau}\right)\right\rangle_0 \quad (64)$$

to first order in $\epsilon = F_e$, for a generic observable $\mathcal{O}$. The function $R(t) = \left\langle J\left(\mathcal{O} \circ S_0^t\right)\right\rangle_0$ is called the response function. In particular, the difference in (56) between exact and linear response can be expressed as:

$$\langle \mathcal{O} \rangle_{t;\epsilon} - \langle \mathcal{O} \rangle_{\widetilde{t;\epsilon}} = \epsilon \int_0^t d\tau \left\langle \Omega_{\mathbf{G}_p}^{f_0}\left[\mathcal{O} \circ S_\epsilon^{\tau} - \mathcal{O} \circ S_0^{t-\tau}\right]\right\rangle_0 \quad (65)$$

That the full dynamics S_ϵ^τ can be replaced by the equilibrium dynamics S_0^τ is a far from trivial fact. In fact, van Kampen raised one of the best known objections to this theory, based on the observation that particles dynamics are typically chaotic [48]. He maintained that while there is no doubt that linear response works in practice, the derivations a la Green and Kubo do not explain why. He stated that: *"Linear response theory does provide expressions for the phenomenological coefficients, but I assert that it arrives at these expressions by a mathematical exercise, rather than by describing the actual mechanism which is responsible for the response"* [48]. It fact, to justify the Green–Kubo calculation, one must invoke non trivial ingredients, such as the large value of N, which makes the perturbation per degree of freedom negligible, and the mingling of contrasting contributions operated by chaos, see e.g., [45,49,50]. As such conditions actually hold in macroscopic particles systems under quite variegated conditions, the Green–Kubo reasoning explains the success of linear response.

3.1. Exact Response and GK Formalism: Example 1

Consider a particle in a two-dimensional space, whose unperturbed Hamiltonian is $H_0 = (p_x^2 + p_y^2)/2$. Take the canonical ensemble $f_0 = Z_0^{-1} \exp(-\beta H_0)$ as the initial distribution. Introduce an external field with Hamiltonian $H_p = -Eq_x$, as a perturbation to obtain $H = H_0 + H_p$. The corresponding equations of motion are:

$$\begin{cases} \dot{q}_x = p_x, \\ \dot{q}_y = p_y, \\ \dot{p}_x = E, \\ \dot{p}_y = 0, \end{cases} \tag{66}$$

with the dissipation function

$$\Omega_E^{f_0} = \Omega_0^{f_0} + \Omega_p^{f_0} = \beta E p_x,$$

in which the momentum in x-direction plays the role of internal dissipative flux with E the corresponding thermodynamic force, i.e., $J = -p_x$ and $F_e = E$. Note that the present example is an instance of the adiabatic system (15), in which $C_1 = C_2 = 0$ and $D_1 = 1, D_2 = 0$, see (59). Sticking to the notation of Section 2.1 in the following we denote by $\Gamma = (q_x, q_y, p_x, p_y)$ the phase space points and observe that the solution of (66) with initial condition Γ gives $p_x(\Gamma, t) = p_x + Et$ and $p_y(\Gamma, t) = p_y$.

Then, thanks to Equation (5), we compute the time evolution of the probability distribution function $f_t(\Gamma)$ starting from the integrated dissipation function:

$$\Omega_{-t,0}^{f_0}(\Gamma) = \beta E \int_{-t}^0 p_x(\Gamma, s) \mathrm{d}s = \beta E p_x t - \beta E^2 \frac{t^2}{2},$$

obtaining

$$f_t(\Gamma) = f_0(\Gamma) e^{\Omega_{-t,0}^{f_0}(\Gamma)} = f_0(\Gamma) \exp\left[\beta E p_x t - \beta E^2 \frac{t^2}{2}\right].$$

Following the discussion in Section 3, we expand the exponential in the limit of small external field E, and we neglect the terms of order higher than linear in E, to obtain the linear response for the evolving density:

$$\tilde{f}_t(\Gamma) = f_0(\Gamma)(1 + \beta E p_x t).$$

The difference between exact and linear response can be further clarified by looking at the Liouville equation for $f_t(\Gamma)$ and proceeding à la Green–Kubo. Let us start introducing the Liouville operator of the perturbed system:

$$\mathcal{L} = \mathcal{L}_0 + \Delta \mathcal{L},$$

where $\mathcal{L}_0 = p_x \partial_{q_x} + p_y \partial_{q_y}$ corresponds to the unperturbed system and $\Delta \mathcal{L} \equiv \mathcal{L}_p = E \partial_{p_x}$ corresponds to the perturbation. Splitting the evolving ensemble distribution function as:

$$f_t(\Gamma) = f_0(\Gamma) + \Delta f_t(\Gamma)$$

the Liouville equation takes the form:

$$\partial_t(f_0(\Gamma) + \Delta f_t(\Gamma)) + (\mathcal{L}_0 + \Delta \mathcal{L})(f_0(\Gamma) + \Delta f_t(\Gamma)) = 0$$

Recalling that the equilibrium distribution $f_0(\Gamma)$ is invariant under the unperturbed dynamics—so that $(\partial_t + \mathcal{L}_0)f_0(\Gamma) = 0$—we obtain the evolution equation for the deviation $\Delta f_t(\Gamma)$:

$$(\partial_t + \mathcal{L}_0)\Delta f_t(\Gamma) = -\Delta \mathcal{L} f_0(\Gamma) - \Delta \mathcal{L} \Delta f_t(\Gamma).$$

Then, assuming that $\Delta \mathcal{L} \Delta f_t(\Gamma)$ is negligible w.r.t. the other terms, we arrive at:

$$(\partial_t + \mathcal{L}_0)\Delta f_t(\Gamma) = -\Delta \mathcal{L} f_0(\Gamma). \tag{67}$$

Since $\Delta \mathcal{L} f_0(\Gamma) = E \partial_{p_x}\left[Z_0^{-1} \exp(-\beta H_0)\right] = -\beta E p_x f_0(\Gamma)$, the Equation (67) can be rewritten as

$$(\partial_t + p_x \partial_{q_x} + p_y \partial_{q_y})\Delta f_t(\Gamma) = \beta E p_x f_0(\Gamma)$$

where $\beta f_0 = -\partial f_0 / \partial H_0$, $E = F_e$ and $p_x = -J$, according to standard notation. The solution to the evolution equation for the variation of the ensemble distribution function, is given by:

$$
\begin{aligned}
\Delta f_t(\Gamma) &= -\int_0^t d\tau \exp[-(t-\tau)\mathcal{L}_0]\Delta \mathcal{L}(\tau)f_0(\Gamma) = -\beta E \int_0^t d\tau \exp[-(t-\tau)\mathcal{L}_0](-p_x)f_0(\Gamma) \\
&= -\beta F_e \int_0^t d\tau \exp[-(t-\tau)\mathcal{L}_0]J(\Gamma)f_0(\Gamma),
\end{aligned}
\tag{68}
$$

where we have used the fact that the external perturbation is not time dependent.

Let us now compute the phase space average of an observable $\mathcal{O}$, using $f_t(\Gamma) = f_0(\Gamma) + \Delta f_t(\Gamma)$. We can write:

$$
\begin{aligned}
\langle \mathcal{O} \rangle_t^{\sim} &= \langle \mathcal{O} \rangle_0 - \beta F_e \int_0^t d\tau \int_{\mathcal{M}} d\Gamma f_0(\Gamma)J(\Gamma)\exp[-(t-\tau)\mathcal{L}_0]\mathcal{O}(\Gamma) \\
&= \langle \mathcal{O} \rangle_0 + \beta E \int_0^t d\tau \int_{\mathcal{M}} d\Gamma f_0(\Gamma)p_x(O \circ S_0^{t-\tau}(\Gamma)) = \langle \mathcal{O} \rangle_0 + \int_0^t d\tau \left\langle (O \circ S_0^{t-\tau}) \cdot \Omega_E^{f_0} \right\rangle_0
\end{aligned}
\tag{69}
$$

noting that S_0^t is the flow of the unperturbed dynamics. As an example, take $\mathcal{O} = p_x$, and compute its time evolving linear response. Since $\langle p_x \rangle_0 = 0$ and $\langle p_x^2 \rangle_0 = \beta^{-1}$, we have:

$$\langle p_x \rangle_t^{\sim} = \frac{\beta E}{Z_0} \int_0^t ds \int dp_y e^{-\frac{\beta}{2}p_y^2} \int dp_x e^{-\frac{\beta}{2}p_x^2} p_x^2 = \beta E \left\langle p_x^2 \right\rangle_0 t = \beta E \frac{1}{\beta} t = Et \tag{70}$$

Now, denoting by S_E^t the flow of the perturbed system, and recalling that $p_x \circ S_E^t = p_x + Et$ and $\Omega_E^{f_0} = \beta E p_x$, the exact response formula (8) yields:

$$\langle p_x \rangle_t = \int_0^t d\tau \int_{\mathcal{M}} d\Gamma f_0(\Gamma)(p_x \circ S_E^\tau(\Gamma))\Omega_E^{f_0}(\Gamma) = \beta E \int_0^t d\tau \int_{\mathcal{M}} d\Gamma f_0(\Gamma)(p_x^2 + Ep_x\tau) = \beta E \left\langle p_x^2 \right\rangle_0 t = Et \tag{71}$$

For such an observable, linear and exact response coincide. In general, we have:

$$\langle \mathcal{O} \rangle_t - \langle \mathcal{O} \rangle_{\tilde{t}} = \int_0^t ds \left\langle \Omega_E^{f_0} [(\mathcal{O} \circ S_E^\tau) - (\mathcal{O} \circ S_0^{t-\tau})] \right\rangle_0. \tag{72}$$

For an analytic observable that do not depend on q_x and q_y, we have:

$$\mathcal{O}(S_E^t \Gamma) = \mathcal{O}(p_x + Et, p_y) = \mathcal{O}(p_x, p_y) + \sum_{k=1}^{+\infty} \frac{\mathcal{O}^{(k)}(p_x, p_y)}{k!} (Et)^k, \tag{73}$$

assuming that $\mathcal{O} = \mathcal{O}(p_x, p_y)$ is an analytic function, and denoting by $\mathcal{O}^{(k)}$ the $k-$th partial derivative of $\mathcal{O}$ with respect to p_x, at fixed p_y. The difference between the exact and the linear response would read:

$$\langle \mathcal{O} \rangle_t - \langle \mathcal{O} \rangle_{\tilde{t}} = \int_0^t d\tau \left\langle \Omega_E^{f_0}(\Gamma) \sum_{k=1}^{+\infty} \frac{\mathcal{O}^{(k)}(p_x, p_y)}{k!} (E\tau)^k \right\rangle_0 \tag{74}$$

Depending on the particular functional form of $\mathcal{O}$ this quantity may grow rapidly or less rapidly with t and, as long as $\mathcal{O}(S_E^t \Gamma)$ exists at all t, this difference at fixed t goes to 0 as $E \to 0$.

3.2. Exact Response and GK Formalism: Example 2

Let the unperturbed Hamiltonian correspond to a harmonic oscillator $H_0 = (p^2 + \omega^2 q^2)/2$, and let the perturbation take the form $H_p = -Eq$. The equations of motion of $H = H_0 + H_p$ are:

$$\begin{cases} \dot{q} = p, \\ \dot{p} = E - \omega^2 q, \end{cases} \tag{75}$$

with flow

$$S_E^t \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} \left(q - \frac{E}{\omega^2}\right) \cos \omega t + \frac{p}{\omega} \sin \omega t + \frac{E}{\omega^2} \\ -\omega \left(q - \frac{E}{\omega^2}\right) \sin \omega t + p \cos \omega t \end{pmatrix}. \tag{76}$$

The dissipation function with initial distribution $f_0 = e^{-\beta H_0}/Z_0$ is given by:

$$\Omega_E^{f_0} = -\mathbf{G} \cdot \nabla \ln f_0 = (p, E - \omega^2 q) \cdot \beta \begin{pmatrix} \omega^2 q \\ p \end{pmatrix} = \beta E p \tag{77}$$

since the vector field is Hamiltonian and Λ vanishes. Moreover, we see that $\Omega_E^{f_0}$ coincides with the dissipation function $\Omega_p^{f_0}$ of the vector filed of H_p, because f_0 is the equilibrium distribution of H_0. Thus the exact and the linear response formulas are respectively given by:

$$\langle \mathcal{O} \rangle_t = \langle \mathcal{O} \rangle_0 + \int_0^t ds \, \langle \beta E p \, (\mathcal{O} \circ S_E^s) \rangle_0 \tag{78}$$

and

$$\langle \mathcal{O} \rangle_{\tilde{t}} = \langle \mathcal{O} \rangle_0 + \int_0^t ds \, \langle \beta E p \, (\mathcal{O} \circ S_0^{t-s}) \rangle_0. \tag{79}$$

Consider now $\mathcal{O} = q$. In this case $\mathcal{O} \circ S_E^s$ and $\mathcal{O} \circ S_0^s$ are given by (76). For $E \neq 0$ we have:

$$\begin{aligned} \langle \beta E p \, (\mathcal{O} \circ S_E^s) \rangle_0 &= \int \int dq \, dp \frac{e^{-\beta(\frac{p^2}{2} + \frac{\omega^2}{2} q^2)}}{Z_0} \beta E p \left[\left(q - \frac{E}{\omega^2}\right) \cos \omega t + \frac{p}{\omega} \sin \omega t + \frac{E}{\omega^2} \right] \\ &= \frac{\beta E}{\omega} \sin \omega t \int_{-\infty}^{+\infty} dq \int_{-\infty}^{+\infty} dp \frac{e^{-\beta(\frac{p^2}{2} + \frac{\omega^2}{2} q^2)}}{Z_0} p^2 = \frac{E}{\omega} \sin \omega t, \end{aligned} \tag{80}$$

since $\langle p \rangle_0 = \int \int dq\, dp\, f_0(q,p)p = 0$, $\langle q \rangle_0 = \int \int dq\, dp\, f_0(q,p)q = 0$ and $\langle p^2 \rangle_0 = \int \int dq\, dp\, f_0(q,p)p^2 = \beta^{-1}$. The response formula (78) gives

$$\langle p \rangle_t = \langle p \rangle_0 + \int_0^t ds\, \frac{E}{\omega} \sin \omega s = \frac{E}{\omega^2}(1 - \cos \omega t). \tag{81}$$

Analogously, the unperturbed dynamics, given by $E = 0$ in (76), yields:

$$\langle \beta E p\, (\mathcal{O} \circ S_0^{t-s}) \rangle_0 = \int \int dq\, dp\, \frac{e^{-\beta(\frac{p^2}{2} + \frac{\omega^2}{2}q^2)}}{Z_0} \beta E p\, \left[q \cos \omega\, (t-s) + \frac{p}{\omega} \sin \omega\, (t-s) \right] = \frac{E}{\omega} \sin \omega\, (t-s), \tag{82}$$

and from (79)

$$\langle p \rangle_{\widetilde{t}} = \langle p \rangle_0 + \int_0^t ds\, \frac{E}{\omega}\, \sin \omega\, (t-s) = \frac{E}{\omega^2}(1 - \cos \omega t). \tag{83}$$

This shows that the perturbed and unperturbed responses of the observable $\mathcal{O} = p$ coincide. If we consider the observable $\mathcal{O} = q\,p$, the situation is different. In the first place,

$$\mathcal{O} \circ S_E^t = -\omega \left(q - \frac{E}{\omega^2} \right)^2 \frac{1}{2} \sin 2\omega t + p \left(q - \frac{E}{\omega^2} \right) \cos^2 \omega t - p \left(q - \frac{E}{\omega^2} \right) \sin^2 \omega t + p^2 \frac{1}{2\omega} \sin 2\omega t$$
$$- \frac{E}{\omega} \left(q - \frac{E}{\omega^2} \right) \sin \omega t + p \frac{E}{\omega^2} \cos \omega t, \tag{84}$$

that must be multiplied by p before taking the ensemble averages in (78) and (79). This gives a linear combination of monomials p, pq, pq^2, p^2q, p^3, whose average with respect to $\langle \cdot \rangle_0$ vanishes, and a non vanishing term

$$p^2 \frac{E}{\omega^2} \left[- \cos^2 \omega t + \sin^2 \omega t + \cos \omega t \right]$$

Thus, recalling that $\langle p^2 \rangle_0 = \beta^{-1}$, we have:

$$\langle qp \rangle_t = \beta E \int_0^t ds\, \frac{E}{\omega^2} \langle p^2 \rangle_0 \left[- \cos^2 \omega s + \sin^2 \omega s + \cos \omega s \right] = \frac{E^2}{\omega^3} \left[\sin \omega t + \frac{1}{2} \sin 2\omega t \right]. \tag{85}$$

On the other hand, from (84) with $E = 0$, we obtain that $p(\mathcal{O} \circ S_0^t)$ is a linear combination of $q^2 p$, qp^2, p^3. This implies that $\langle \beta E p\, (\mathcal{O} \circ S_0^{t-s}) \rangle_0 = 0$ and, from (79), that $\langle qp \rangle_{\widetilde{t}} = 0$, showing that the perturbed and unperturbed responses substantially differ in the case of $\mathcal{O} = q\,p$, unless E is particularly small.

4. Formal Thermodynamics

Following Qian and collaborators [29], we now obtain various new relations for the dissipation function. To this end, suppose that the system admits an energy function $H : \mathcal{M} \to \mathbb{R}$, $H \in C^1(\mathbb{R}^N, \mathbb{R})$. Let us introduce the set Γ_h in phase space, defined as follows:

$$\Gamma_h = \{ \mathbf{x} \in \mathbb{R}^N : H(\mathbf{x}) \leq h \}$$

and its boundary:

$$\partial \Gamma_h = \{ \mathbf{x} \in \mathbb{R}^N : H(\mathbf{x}) = h \}$$

The set Γ_h includes all points in phase space whose energy is equal to or lower than the threshold h, and $\partial \Gamma_h$ is its boundary, constituted by the phase points whose energy equals the threshold value. Introduce the measure of such sets w.r.t. the time evolving probability measure:

$$\mu_t(h) = \int_{\Gamma_h} f_t(\mathbf{x})\, d\mathbf{x}$$

where $f_t(\mathbf{x})$ is a sufficiently well behaved probability density. Then, the function

$$\frac{f_t(\mathbf{x})}{\mu_t(h)}\chi_h(x); \qquad \chi_h = \begin{cases} 1 & \text{if } x \in \Gamma_h \\ 0 & \text{if } x \notin \Gamma_h \end{cases} \tag{86}$$

where χ_h is the characteristic function of Γ_h, plays the role of a probability density in Γ_h. Using Reynolds Transport Theorem plus the regularity of the surface $\partial\Gamma_h$, one finds:

$$\frac{d}{dh}\mu_t(h) = \mu'_t(h) = \int_{\partial\Gamma_h} f_t(\mathbf{x})\frac{d\sigma}{\|\nabla H\|}, \tag{87}$$

where $d\sigma$ is the surface element area. This quantity can be interpreted as the measure w.r.t. the density $f_t(\mathbf{x})$ of the constant energy surface $\partial\Gamma_h$. The time derivative along a trajectory of the energy function $H(\mathbf{x})$ is given by:

$$\dot{H}(\mathbf{x}) = \mathbf{V}(\mathbf{x})\cdot\nabla H(\mathbf{x})$$

hence we may define a formal heat flux through the boundary $\partial\Gamma_h$ as:

$$\dot{Q}_t(h) = -\frac{1}{\mu'_t(h)}\int_{\partial\Gamma_h}\mathbf{V}(\mathbf{x})\cdot\nabla H(\mathbf{x})f_t(\mathbf{x})\frac{d\sigma}{\|\nabla H\|}$$

If in addition, we introduce the formal temperature as the inverse of the logarithmic derivative of the constant energy volume w.r.t. the energy parameter h:

$$T_t(h) = \left(\frac{d}{dh}\ln\mu_t(h)\right)^{-1} = \frac{\mu_t(h)}{\mu'_t(h)}$$

we can relate the ensuing formal entropy production rate to the average of the dissipation function concerning the set Γ_h:

$$e_p(h;t) = \frac{\dot{Q}_t(h)}{T_t(h)} = \frac{1}{\mu_t(h)}\int_{\Gamma_h}\Omega^{f_t}(\mathbf{x})f_t(\mathbf{x})\,d\mathbf{x} = \frac{1}{\mu_t(h)}\left\langle\Omega^{f_t}\cdot\chi_h\right\rangle_t \tag{88}$$

This result is readily obtained as follows:

$$\dot{Q}_t(h) = -\frac{1}{\mu'_t(h)}\int_{\partial\Gamma_h}\mathbf{V}(\mathbf{x})\cdot\nabla H(\mathbf{x})f_t(\mathbf{x})\frac{d\sigma}{\|\nabla H\|} = -\frac{1}{\mu'_t(h)}\int_{\Gamma_h}\nabla\cdot(f_t(\mathbf{x})\mathbf{V}(\mathbf{x}))\,d\mathbf{x}$$

where the equality is guaranteed by Divergence Theorem, $\nabla H/\|\nabla H\|$ being the unit vector everywhere normal to the surface $\partial\Gamma_h$, and by the identity $\nabla\cdot(f_t(\mathbf{x})\mathbf{V}(\mathbf{x})) = -\Omega^{f_t}(\mathbf{x})f_t(\mathbf{x})$. Then Equations (3) and (5), yield:

$$\Omega^{f_t}(\mathbf{x}) = -(\Lambda(\mathbf{x}) + \mathbf{V}(\mathbf{x})\cdot\nabla\ln f_t(\mathbf{x})) = -\left[\Lambda(\mathbf{x}) + \mathbf{V}(\mathbf{x})\cdot\nabla\ln(f_0(\mathbf{x})\exp\{\Omega^{f_0}_{-t,0}(\mathbf{x})\}\right] =$$
$$= -\left[\Lambda(\mathbf{x}) + \mathbf{V}(\mathbf{x})\cdot\nabla(\ln f_0(\mathbf{x}) + \Omega^{f_0}_{-t,0}(\mathbf{x}))\right] \tag{89}$$

that leads to:

$$\Omega^{f_t}(\mathbf{x}) = \Omega^{f_0}(\mathbf{x}) - \mathbf{V}(\mathbf{x})\cdot\nabla\Omega^{f_0}_{-t,0}(\mathbf{x}). \tag{90}$$

Therefore, we can now write:

$$e_p(h;t) = \frac{1}{\mu_t(h)}\int_{\mathcal{M}}\Omega^{f_t}(\mathbf{x})\chi_h(\mathbf{x})f_t(\mathbf{x})\,d\mathbf{x} = \frac{1}{\mu_t(h)}\int_{\mathcal{M}}\left(\Omega^{f_0}(\mathbf{x}) - \mathbf{V}(\mathbf{x})\cdot\nabla\Omega^{f_0}_{-t,0}(\mathbf{x})\right)\chi_h(\mathbf{x})f_t(\mathbf{x})\,d\mathbf{x}$$
$$= \frac{1}{\mu_t(h)}\left(\int_{\mathcal{M}}\Omega^{f_0}(\mathbf{x})\chi_h(\mathbf{x})f_t(\mathbf{x})\,d\mathbf{x} - \int_{\mathcal{M}}\mathbf{V}(\mathbf{x})\cdot\nabla\Omega^{f_0}_{-t,0}(\mathbf{x})\chi_h(\mathbf{x})f_t(\mathbf{x})\,d\mathbf{x}\right) \tag{91}$$

The first addend within the brackets can be reworked using the response formula (8), where Ω^{f_0} restricted to Γ_h is considered like any other phase function:

$$\int_{\mathcal{M}} \Omega^{f_0}(\mathbf{x})\chi_h(\mathbf{x})f_t(\mathbf{x})\,\mathrm{d}x = \int_{\mathcal{M}} \Omega^{f_0}(\mathbf{x})\chi_h(\mathbf{x})f_0(\mathbf{x})\,\mathrm{d}x + \int_0^t \int_{\mathcal{M}} \Omega^{f_0}(S^\tau\mathbf{x})\chi_h(S^\tau\mathbf{x})\Omega^{f_0}(\mathbf{x})f_0(\mathbf{x})\,\mathrm{d}x\,\mathrm{d}\tau \quad (92)$$

while for the second contribution, we have:

$$
\begin{aligned}
\int_{\mathcal{M}} \mathbf{V}(\mathbf{x}) \cdot \nabla\Omega^{f_0}_{-t,0}(\mathbf{x})\chi_h(\mathbf{x})f_t(\mathbf{x})\,\mathrm{d}x &= \int_{\mathcal{M}} \mathbf{V}(\mathbf{x}) \cdot \nabla\Omega^{f_0}_{-t,0}(\mathbf{x})\chi_h(\mathbf{x})f_0(\mathbf{x})e^{\Omega^{f_0}_{-t,0}(\mathbf{x})}\,\mathrm{d}x \\
&= \int_{\mathcal{M}} \mathbf{V}(\mathbf{x}) \cdot \nabla e^{\Omega^{f_0}_{-t,0}(\mathbf{x})}\chi_h(\mathbf{x})f_0(\mathbf{x})\,\mathrm{d}x
\end{aligned}
\quad (93)
$$

Rearranging the two contributions, the first integral in Equation (91) becomes:

$$\left\langle \Omega^{f_t}\chi_h \right\rangle_t = \left\langle \Omega^{f_0}\chi_h \right\rangle_0 + \int_0^t \mathrm{d}\tau \int_{\mathcal{M}} \mathrm{d}x\, \Omega^{f_0}(S^\tau\mathbf{x})\chi_h(S^\tau\mathbf{x})\Omega^{f_0}(\mathbf{x})f_0(\mathbf{x}) - \left\langle \mathbf{V}\cdot\nabla e^{\Omega^{f_0}_{-t,0}}\chi_h \right\rangle_0 \quad (94)$$

and finally:

$$e_p(h;t) = \frac{1}{\mu_t(h)}\left\{ \left\langle \left(\Omega^{f_0} - \mathbf{V}\cdot\nabla e^{\Omega^{f_0}_{-t,0}}\right)\chi_h \right\rangle_0 + \int_0^t \left\langle \Omega^{f_0}\left(\Omega^{f_0}\chi_h\right)\circ S^\tau \right\rangle_0 \mathrm{d}\tau \right\}. \quad (95)$$

Letting h grow in Equation (88), these conditional averages converge to the average on $\mathcal{M}$, yielding:

$$\lim_{h\to\infty} e_p(h;t) = \langle\Omega^{f_t}\rangle_t = 0$$

where (13) has been used. This is consistent with the fact that in this model no "energy" can flow out of $\mathcal{M}$, or in $\mathcal{M}$ from outside.

Note that such a reasoning actually does not refer to a physical flow: in the first place, it occurs in phase space, rather than in real space; furthermore the dynamics does not need to describe any physical phenomenon. Therefore, the thermodynamic interpretation of the above relations should be taken with a grain of salt [43,44], while it may be suggestive in the analysis of dynamical systems, analogously to the thermodynamic formalism [33].

The Case with Constant or Slowly-Varying Volume Contraction

Assume that the probability density f_0 is continuous with compact support $\mathcal{M}$, and take $\mathbf{V}(\mathbf{x}) = \mathbf{V}_0(\mathbf{x}) + \epsilon\mathbf{V}_1(\mathbf{x})$, with $\mathbf{V}_0$, $\mathbf{V}_1$ sufficiently regular, $\Lambda_0(\mathbf{x}) \equiv \mathrm{div}\mathbf{V}_0(\mathbf{x}) = -\lambda_0 < 0$. Assume also that the system is $\Omega\,T$-mixing. To compute $\left\langle\Omega_V^{f_0}\right\rangle_t$ up to $O(\epsilon)$, observe that:

$$\left\langle\Omega_V^{f_0}\right\rangle_t = \int_0^t \left\langle\Omega_V^{f_0}\left(\Omega_V^{f_0}\circ S_\epsilon^\tau\right)\right\rangle_0 \mathrm{d}\tau = \int_0^t \left\langle\Omega_V^{f_0}\left(\Omega_{V_0}^{f_0}\circ S_\epsilon^\tau\right)\right\rangle_0 \mathrm{d}\tau + \epsilon\int_0^t \left\langle\Omega_V^{f_0}\left(\Omega_{V_1}^{f_0}\circ S_\epsilon^\tau\right)\right\rangle_0 \mathrm{d}\tau \quad (96)$$

where

$$\Omega_V^{f_0}(\mathbf{x}) = \Omega_{V_0}^{f_0}(\mathbf{x}) + \epsilon\,\Omega_{V_1}^{f_0}(\mathbf{x}) = -\lambda_0 - \mathbf{V}_0(\mathbf{x})\cdot\nabla\ln f_0(\mathbf{x}) + \epsilon\,\Omega_{V_1}^{f_0}(\mathbf{x})$$

Substituting in (96), we have:

$$
\begin{aligned}
\left\langle\Omega_V^{f_0}\right\rangle_t = &\ \lambda_0^2\, t \\
&+ \lambda_0 \int_0^t \mathrm{d}\tau \int_{\mathcal{M}} \mathrm{d}\mathbf{x}\, f_0(\mathbf{x})\mathbf{V}_0(S_\epsilon^\tau\mathbf{x})\cdot\nabla\ln f_0(S_\epsilon^\tau\mathbf{x}) + \lambda_0 \int_0^t \mathrm{d}\tau \int_{\mathcal{M}} \mathrm{d}\mathbf{x}\, f_0(\mathbf{x})\mathbf{V}_0(\mathbf{x})\cdot\nabla\ln f_0(\mathbf{x}) \\
&+ \int_0^t \mathrm{d}\tau \int \mathrm{d}\mathbf{x}\, f_0(\mathbf{x})(\mathbf{V}_0(S_\epsilon^\tau\mathbf{x})\cdot\nabla\ln f_0(S_\epsilon^\tau\mathbf{x}))\,(\mathbf{V}_0(\mathbf{x})\cdot\nabla\ln f_0(\mathbf{x})) + O(\epsilon)
\end{aligned}
\quad (97)
$$

The second integral can be computed explicitly taking $f_0(\mathbf{x}) = 0$ on $\partial \mathcal{M}$:

$$\langle \mathbf{V}_0 \cdot \nabla \ln f_0 \rangle_0 = \int_{\mathcal{M}} \mathbf{V}_0(\mathbf{x}) \cdot \nabla f_0(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{M}} \mathrm{div}(\mathbf{V}_0(\mathbf{x}) f_0(\mathbf{x})) d\mathbf{x} - \int_{\mathcal{M}} \mathrm{div} \mathbf{V}_0(\mathbf{x}) f_0(\mathbf{x}) d\mathbf{x}. \tag{98}$$

As the divergence of $\mathbf{V}_0$ equals the constant $-\lambda_0$, and $\int_{\mathcal{M}} f_0(\mathbf{x}) d\mathbf{x} = 1$, we have:

$$\langle \mathbf{V}_0 \cdot \nabla \ln f_0 \rangle_0 = \lambda_0 + \int_{\mathcal{M}} \mathrm{div}(\mathbf{V}_0(\mathbf{x}) f_0(\mathbf{x})) d\mathbf{x}. \tag{99}$$

The integral in the previous equation is, by the Divergence Theorem, equal to

$$\int_{\partial M} f_0(\mathbf{x}) V_0(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) d\sigma = 0,$$

Then

$$\langle \mathbf{V}_0 \cdot \nabla \ln f_0 \rangle_0 = \lambda_0, \tag{100}$$

and

$$\left\langle \Omega_{\mathbf{V}}^{f_0} \right\rangle_t = 2\lambda_0^2 t + \lambda_0 \int_0^t \langle \mathbf{V}_0(S_\epsilon^\tau(\mathbf{x})) \cdot \nabla \ln f_0(S_\epsilon^\tau(\mathbf{x})) \rangle_0 d\tau + \\
+ \int_0^t \langle (\mathbf{V}_0(S_\epsilon^\tau(\mathbf{x})) \cdot \nabla \ln f_0(S_\epsilon^\tau(\mathbf{x}))) (\mathbf{V}_0(\mathbf{x}) \cdot \nabla \ln f_0(\mathbf{x})) \rangle_0 d\tau + O(\epsilon). \tag{101}$$

Averaging Λ one has instead:

$$\langle \Lambda \rangle_t = \int_{\mathcal{M}} \Lambda(\mathbf{x}) f_t(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{M}} (\lambda_0 - \epsilon \, \mathrm{div} \mathbf{V}_1(\mathbf{x})) \, f_t(\mathbf{x}) d\mathbf{x} = \lambda_0 + O(\epsilon),$$

since to first order, our vector field has a constant phase space volumes contraction rate.

Example 1. *The simplest case is obtained when the initial distribution is uniform, i.e., $f_0(\mathbf{x}) =$ constant, as if the "equilibrium" vector field vanishes. In that case the integrals in* (101) *vanish, and*

$$\left\langle \Omega_{\mathbf{V}}^{f_0} \right\rangle_t = 2\lambda_0^2 t + O(\epsilon) \,.$$

Example 2. *Suppose the coordinates in the initial ensemble are independent, i.e.,*

$$f_0(\mathbf{x}) = \prod_{i=1}^n g_i(x_i),$$

with g_i being one dimensional density functions. Then,

$$\frac{\partial}{\partial x_k} \ln f_0(\mathbf{x}) = \frac{g_k'(x_k)}{g_k(x_k)} \tag{102}$$

and Equation (101) *can be rewritten as follows:*

$$\left\langle \Omega_{\mathbf{V}}^{f_0} \right\rangle_t = 2\lambda_0^2 t + \lambda_0 \sum_{i=1}^n \int_0^t \left\langle V_{0i}(S_\epsilon^\tau \mathbf{x}) \frac{g_i'((S_\epsilon^\tau \mathbf{x})_i)}{g_i((S_\epsilon^\tau \mathbf{x})_i)} \right\rangle_0 d\tau + \\
+ \sum_{i,j=1}^n \int_0^t \left\langle V_{0i}(S_\epsilon^\tau \mathbf{x}) V_{0j}(\mathbf{x}) \frac{g_i'((S_\epsilon^\tau \mathbf{x})_i)}{g_i((S_\epsilon^\tau \mathbf{x})_i)} \frac{g_j'(x_j)}{g_j(x_j)} \right\rangle_0 d\tau + O(\epsilon). \tag{103}$$

In the special case of Gaussian marginals

$$g_k(x_k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(x_k - \mu_k)^2}{2\sigma_k^2}} \; ; \qquad \frac{g_i'(x_i)}{g_i(x_i)} = -\frac{(x_k - \mu_k)}{\sigma_k^2} \tag{104}$$

we have:

$$\left\langle \Omega_{\mathbf{V}}^{f_0} \right\rangle_t = 2\lambda_0^2 t - \lambda_0 \sum_{i=1}^n \sigma_i^{-2} \int_0^t \left\langle V_{0i}(S_\epsilon^\tau \mathbf{x})((S_\epsilon^\tau \mathbf{x})_i - \mu_i) \right\rangle_0 d\tau + \tag{105}$$
$$+ \sum_{i,j=1}^n (\sigma_i \sigma_j)^{-2} \int_0^t \left\langle V_{0i}(S_\epsilon^\tau \mathbf{x}) V_{0j}(\mathbf{x})((S_\epsilon^\tau \mathbf{x})_i - \mu_i)(x_j - \mu_j) \right\rangle_0 d\tau + O(\epsilon).$$

The simplest example can be obtained with $\epsilon = 0$ and $V_{0i}(\mathbf{x}) = -\lambda_i x_i$, $\lambda_i > 0$, and assuming $\mu_k = 0$. Then, observing that the i-th component of $S_0^\tau \mathbf{x}$, can be written as $(S_0^\tau \mathbf{x})_i = e^{-\lambda_i t} x_i$, Equation (105) becomes:

$$\left\langle \Omega_{\mathbf{V}}^{f_0} \right\rangle_t = 2\lambda_0^2 t + \lambda_0 \sum_{i=1}^n \sigma_i^{-2} \lambda_i \int_0^t e^{-2\lambda_i \tau} \left\langle x_i^2 \right\rangle_0 d\tau + \tag{106}$$
$$+ \sum_{i,j=1}^n (\sigma_i \sigma_j)^{-2} \lambda_i \lambda_j \int_0^t e^{-2(\lambda_i + \lambda_j)\tau} \left\langle x_i^2 x_j^2 \right\rangle_0 d\tau,$$

where $\lambda_0 = \sum_{i=1}^n \lambda_i$. In addition, recalling that $\left\langle x_i^2 \right\rangle_0 = \sigma_i^2$ and $\left\langle x_i^2 x_j^2 \right\rangle_0 = \sigma_i^2 \sigma_j^2$, we have:

$$\left\langle \Omega_{\mathbf{V}}^{f_0} \right\rangle_t = 2\lambda_0^2 t + \frac{1}{2}\lambda_0 n - \frac{1}{2}\lambda_0 \sum_{i=1}^n e^{-2\lambda_i t} + \sum_{i,j=1}^n \frac{\lambda_i \lambda_j}{\lambda_i + \lambda_j} - \sum_{i,j=1}^n \frac{\lambda_i \lambda_j}{\lambda_i + \lambda_j} e^{-2(\lambda_i + \lambda_j)t}. \tag{107}$$

The previous result could have been obtained directly from Equation (96), noting that:

$$\Omega_{V_0}^{f_0}(\mathbf{x}) = -\lambda_0 - \sum_{i=1}^n \frac{\lambda_i}{\sigma_i^2} x_i^2.$$

Equation (107) shows that the average dissipation function $\left\langle \Omega_{\mathbf{V}}^{f_0} \right\rangle_t$ diverges as $t \to \infty$, because the dynamics has the global attractor $\mathbf{x} = \mathbf{0}$. Thus, $\Omega_{\mathbf{V}}^{f_0}$ tends to the non vanishing phase space contraction rate $-\lambda_0$. That is not the case for $\epsilon > 0$.

5. Simple Attractors and Small Dissipation

Consider the dynamical system (1) on a compact phase space $\mathcal{M}$ containing a set of attracting fixed points $\mathbf{A} = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ with (open) basins of attraction $\mathcal{B}_j = \{\mathbf{x} \in \mathcal{M} : S^t \mathbf{x} \to \mathbf{x}_j, \; t \to +\infty\}$, $j = 1, \ldots, k$, such that $\cup_{i=1}^k \mathcal{B}_j = \mathcal{M}$, modulo a set of zero Lebesgue measure. Let f_0 be the initial probability density and f_t its evolution at time t. Recalling that the basins of attraction are pairwise disjoint, for any continuous phase function C we can write:

$$\langle C \rangle_t = \int_{\mathcal{M}} C(\mathbf{x}) f_t(\mathbf{x}) \, d\mathbf{x} = \int_{\mathcal{M}} C(S^t \mathbf{x}) f_0(\mathbf{x}) \, d\mathbf{x} = \sum_{i=1}^k \int_{\mathcal{B}_i} C(S^t \mathbf{x}) f_0(\mathbf{x}) \, d\mathbf{x}. \tag{108}$$

Taking the limit, by bounded convergence we have:

$$\lim_{t \to \infty} \langle C \rangle_t = \sum_{i=1}^k w_i C(\mathbf{x}_i) \; ; \qquad \text{where} \quad \sum_{i=1}^k w_i = 1, \quad w_i \equiv \int_{\mathcal{B}_i} f_0(\mathbf{x}) \, d\mathbf{x} > 0 \tag{109}$$

This shows that the asymptotic mean of C is the weighted mean of the values of C on the attracting fixed points, with weights given by the μ_0-probability measure of the corresponding basins of attraction.

In other words, the probability measure $d\mu_t(\mathbf{x}) = f_t(\mathbf{x})\,d\mathbf{x}$ converges weakly in time to $\nu = \sum_{i=1}^k w_i \delta_{\mathbf{x}_i}$, with $\delta_{\mathbf{x}_i}$ the atomic measure on $\mathbf{x}_i$. In this case it is easy to compute the correlation function of A and C w.r.t. μ_0:

$$\langle (A \circ S^t) C \rangle_0 - \langle (A \circ S^t) \rangle_0 \langle C \rangle_0 \tag{110}$$

that is used to define T-mixing and, taking $C = \Omega^{f_0}$, to define ΩT-mixing, cf. Equations (10) and (11). The $t \to \infty$ limit of (110) then yields:

$$\lim_{t \to \infty} \left[\langle (A \circ S^t) C \rangle_0 - \langle (A \circ S^t) \rangle_0 \langle C \rangle_0 \right] = \sum_{i=1}^k A(\mathbf{x}_i) \left[\int_{\mathcal{B}_i} C(\mathbf{x}) f_0(\mathbf{x})\,d\mathbf{x} - w_i \int_{\mathcal{M}} C(\mathbf{x}) f_0(\mathbf{x})\,d\mathbf{x} \right] \tag{111}$$

and, assuming that also Ω^{f_0} is continuous,

$$\lim_{t \to \infty} \left[\langle (C \circ S^t) \Omega^{f_0} \rangle_0 - \langle (C \circ S^t) \rangle_0 \langle \Omega^{f_0} \rangle_0 \right] = \lim_{t \to \infty} \langle (C \circ S^t) \Omega^{f_0} \rangle_0 = \sum_{i=1}^k C(\mathbf{x}_i) \int_{\mathcal{B}_i} \Omega^{f_0}(\mathbf{x}) f_0(\mathbf{x})\,d\mathbf{x} \tag{112}$$

since $\langle \Omega^{f_0} \rangle_0 = 0$.

Remark 1. *For $k > 1$, the response from a single initial condition is not unique and depends on the initial distribution f_0. Hence, T-mixing does not hold, since it requires the long time averages to be independent of the initial condition [26]. On the other hand, ΩT-mixing holds, since it is necessary and sufficient for the convergence of the full phase space averages to their asymptotic limit, i.e., (109).*

Now, consider a case with small dissipation:

$$\dot{\mathbf{x}} = \mathbf{V}(\mathbf{x}; \epsilon) = \mathbf{V}_0(\mathbf{x}) + \epsilon \mathbf{V}_1(\mathbf{x}), \tag{113}$$

where ϵ is small, and

$$\mathrm{div}\,\mathbf{V}_0(\mathbf{x}) = 0; \qquad \mathbf{V}_0(\mathbf{x}) \cdot \nabla f_0(\mathbf{x}) = 0 \tag{114}$$

Then, the unperturbed dynamics $\dot{\mathbf{x}} = \mathbf{V}_0(\mathbf{x})$ preserves the volume and the initial density, since $\Omega_{\mathbf{V}_0}^{f_0}(\mathbf{x}) \equiv 0$, while the dissipation function only depends on the perturbation $\epsilon \mathbf{V}_1$:

$$\Omega_{\mathbf{V}}^{f_0}(\mathbf{x}) = \epsilon\, \Omega_{\mathbf{V}_1}^{f_0}(\mathbf{x}) = -\epsilon \left[\mathrm{div}\,\mathbf{V}_1(\mathbf{x}) + \mathbf{V}_1(\mathbf{x}) \cdot \nabla \ln f_0(\mathbf{x}) \right] \tag{115}$$

Let the initial probability density f_0 be positive for every $\mathbf{x}$ in the phase space $\mathcal{M}$, a condition at times called ergodic consistency [30]. In this case, the probability density at time t is given by:

$$f_t(\mathbf{x}) = f_0(\mathbf{x}) \exp \left[\epsilon \int_{-t}^0 \Omega_{\mathbf{V}_1}^{f_0}(S_\epsilon^\tau \mathbf{x})\,d\tau \right],$$

where S_ϵ^τ is the flow of (113). Taking $\alpha \in (0,1)$ and $t_\epsilon = 1/\epsilon^\alpha$, it can be shown that the difference between the evolved density and the initial one is $O(\epsilon^{1-\alpha})$, in the ϵ-dependent time interval bound by $-t_\epsilon$, provided $\Omega_{\mathbf{V}}^{f_0}$ is continuous in $\mathcal{M}$. In particular, one has:

$$\left| \int_{-t_\epsilon}^0 \Omega_{\mathbf{V}_1}^{f_0}(S_\epsilon^\tau \mathbf{x})\,dt_1 \right| \leq \max_{\mathbf{x} \in \mathcal{M}} \left| \Omega_{\mathbf{V}_1}^{f_0}(\mathbf{x}) \right| t_\epsilon = \tilde{\Omega}_{\mathbf{V}_1}^{f_0} t_\epsilon \tag{116}$$

which defines the maximum magnitude of $\Omega_{\mathbf{V}_1}^{f_0}$ in $\mathcal{M}$. Consequently, for $0 < t \leq t_\epsilon$:

$$|f_t(\mathbf{x}) - f_0(\mathbf{x})| = f_0(\mathbf{x}) \left| \exp\left(\epsilon \int_{-t}^0 \Omega_{\mathbf{V}_1}^{f_0}(S_\epsilon^\tau \mathbf{x})\,d\tau \right) - 1 \right| \leq f_0(\mathbf{x}) \left[\exp\left(\tilde{\Omega}_{\mathbf{V}_1}^{f_0} \epsilon^{1-\alpha} \right) - 1 \right] = f_0(\mathbf{x})\, O(\epsilon^{1-\alpha}) \tag{117}$$

We can use this inequality to bound from above the variation of an observable of the system at hand. For a continuous function B, one obtains:

$$|\langle B \rangle_t - \langle B \rangle_0| \leq \int_{\mathcal{M}} |B(\mathbf{x})| \left(\exp\left[\widetilde{\Omega}_{V_1}^{f_0} \epsilon^{1-\alpha} \right] - 1 \right) f_0(\mathbf{x}) \, d\mathbf{x} \leq O(\epsilon^{1-\alpha}) \max_{\mathbf{x} \in \mathcal{M}} |B(\mathbf{x})| \qquad (118)$$

6. Conditions on Dynamics and Probability Density

As observed in [27], the dissipation function can be obtained starting from the Liouville equation:

$$\frac{\partial f_t}{\partial t}(\mathbf{x}) = -f_t(\mathbf{x})\Lambda(\mathbf{x}) - V(\mathbf{x}) \cdot \nabla_{\mathbf{x}} f_t(\mathbf{x}) \qquad (119)$$

and factorizing $f_t(\mathbf{x})$ in the right hand side. The result is Equation (2) with $\Omega^{f_t}(\mathbf{x})$ defined as in (3). The standard conditions of physical interest, under which this operation can be performed are the following:

- the vector field should be everywhere differentiable in the phase space $\mathcal{M}$ for Λ to exist;
- the initial density f_0 should be everywhere positive in $\mathcal{M}$ for its logarithm to exist;
- the initial density f_0 should be everywhere differentiable in $\mathcal{M}$ for its gradient to exist.

This also guarantees that the dissipation function be continuous, which is what is used to derive Equation (8) from Equation (7), via Equation (5). In the physical literature, these requirements are satisfied if (a) the condition sometimes called ergodic consistency holds [30]—f_0 should be the equilibrium ensemble for the dynamics subjected to the same constraints of the nonequilbrium dynamics—and if (b) the particles interactions are sufficiently regular. The effects of discontinuities can be controlled if the dynamics are piecewise regular (as in billiards), or the probability decays sufficiently fast when approaching singularities. A full mathematical treatment of these aspects is far from trivial for physically relevant particle systems, and the success of the theory can only be judged a posteriori, on a case by case basis.

At the same time, various mathematical aspects can be understood considering simple examples, that can be explicitly solved. Let us focus on simple cases in which the above conditions are violated. This is important both to delineate the range of applicability of the theory based on the dissipation function and, at the same time, to find corrections where needed.

When $f_t(\mathbf{x})$ is not differentiable, one may replace (119) by its weak form, i.e.,

$$\int_{\mathcal{M}} h(\mathbf{x}) f_t(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{M}} h(\mathbf{x}) f_0(\mathbf{x}) d\mathbf{x} + \int_0^t \int_{\mathcal{M}} V(\mathbf{x}) \cdot \nabla h(\mathbf{x}) f_\tau(\mathbf{x}) d\mathbf{x} d\tau, \quad \text{for all } h(\mathbf{x}) \in C_0^\infty(\mathcal{M}), \qquad (120)$$

$C_0^\infty(\mathcal{M})$ being the set of smooth functions with compact support in $\mathcal{M}$. The non-differentiability, of f_t, however, is not the only issue. In what follows we take f_0 and f_t with compact support K_0 and K_t in $\mathcal{M}$, while $\mathcal{M}$ is not necessarily compact.

Examples

Example 3. *Consider the dynamical system*

$$\dot{x} = -x; \quad x \in \mathbb{R} \qquad (121)$$

and the following initial density:

$$f_0(x) = \begin{cases} 1/2, & x \in [-1, 1] \\ 0, & else \end{cases} \qquad (122)$$

that is compactly supported and non differentiable at 1 and −1. As the initial condition x evolves like $S^t x = e^{-t}x$, f_t is given by

$$f_t(x) = \frac{1}{2} e^t \chi_{[-e^{-t}, e^{-t}]}(x), \quad \text{with} \quad \chi_{[-e^{-t}, e^{-t}]}(x) = \begin{cases} 1 & \text{if } x \in [-e^{-t}, e^{-t}] \\ 0 & \text{if } x \notin [-e^{-t}, e^{-t}] \end{cases} \tag{123}$$

Then, Equation (119) does not apply, while Equation (120) does. Indeed, the left hand side and the first term in the right hand side of (120) are

$$A := \int_{\mathcal{M}} h(x) f_t(x) dx = \frac{e^t}{2} \int_{-e^{-t}}^{e^{-t}} h(x) dx, \quad \text{and} \quad B := \int_{\mathcal{M}} h(x) f_0(x) dx = \frac{1}{2} \int_{-1}^{1} h(x) dx. \tag{124}$$

The phase integral in the right hand side of (120) yields:

$$\int_{\mathcal{M}} V(\mathbf{x}) \cdot \nabla h(\mathbf{x}) f_s(\mathbf{x}) dx = -\frac{e^s}{2} \int_{-e^{-s}}^{e^{-s}} x h'(x) dx = -\frac{1}{2} \left[h(e^{-s}) + h(-e^{-s}) \right] + \frac{e^s}{2} \int_{-e^{-s}}^{e^{-s}} h(x) dx, \tag{125}$$

and the corresponding time integral becomes:

$$C := \int_0^t ds \int_{\mathcal{M}} dx V(\mathbf{x}) \cdot \nabla h(\mathbf{x}) f_s(\mathbf{x}) = -\frac{1}{2} \int_0^t \left[h(e^{-s}) + h(-e^{-s}) \right] ds + \int_0^t ds \left[\frac{e^s}{2} \int_{-e^{-s}}^{e^{-s}} h(x) dx \right], \tag{126}$$

Then, integrating by parts and using Leibniz's formula in the last integral, we get:

$$\int_0^t ds \left[\frac{e^s}{2} \int_{-e^{-s}}^{e^{-s}} h(x) dx \right] = \frac{e^t}{2} \int_{-e^{-t}}^{e^{-t}} h(x) dx - \frac{1}{2} \int_{-1}^{1} h(x) dx + \frac{1}{2} \int_0^t \left[h(e^{-s}) + h(-e^{-s}) \right] ds, \tag{127}$$

and

$$C = \frac{e^t}{2} \int_{-e^{-t}}^{e^{-t}} h(x) dx - \frac{1}{2} \int_{-1}^{1} h(x) dx. \tag{128}$$

Form (124) we get $A = B + C$, that proves (120).

The probability with density $f_t(x)$ converges weakly to the delta function at $x = 0$, since for any continuous $\mathcal{O}$, the mean value theorem yields:

$$\langle \mathcal{O} \rangle_t = \int \mathcal{O}(x) f_t(x) dx = \frac{e^t}{2} \int_{-e^{-t}}^{e^{-t}} \mathcal{O}(x) dx = \mathcal{O}(x_t) \to \mathcal{O}(0) \quad \text{as} \quad t \to +\infty,$$

for one $x_t \in (-e^{-t}, e^{-t})$. Can this result be obtained via the response formula (8)? Because $\Lambda(x) = -1$, and $df_0/dx = 0$, the dissipation function takes the form:

$$\Omega^{f_0}(x) = -\Lambda(x) - V(x) \frac{f_0'(x)}{f_0(x)} = 1 \tag{129}$$

on the interval $[-1, 1]$. Hence, unlike the cases addressed so far, $\left\langle \Omega^{f_0} \right\rangle_0$ does not vanish, it equals 1!

More generally, given the 1-D system $\dot{x} = V(x)$ with $x \in \mathbb{R}$ and f_0 supported on any interval $[a, b]$, integration by parts yields:

$$\left\langle \Omega^{f_0} \right\rangle_0 = \int_a^b \left(-V'(x) - V(x) \frac{f_0'(x)}{f_0(x)} \right) f_0(x) dx = V(a) f_0(a) - V(b) f_0(b), \tag{130}$$

showing that $\left\langle \Omega^{f_0} \right\rangle_0$ vanishes only if some conditions, such as $V(a) = V(b) = 0$ or $f_0(a) = f_0(b) = 0$, are verified.

In Ref. [27], the equality $\left\langle \Omega^{f_t} \right\rangle_{f_t} = 0$ was obtained from the conservation of probability and from the interchange of the time derivative with the integration in phase space:

$$\left\langle \Omega^{f_t} \right\rangle_{f_t} = \int_{\mathcal{M}} \Omega^{f_t} f_t d\Gamma = \int_{\mathcal{M}} \frac{\partial f_t}{\partial t} d\Gamma = \frac{d}{dt} \int_{\mathcal{M}} f_t d\Gamma = \frac{d}{dt} 1 = 0 \tag{131}$$

This exchange is legitimate under the joint continuity w.r.t. x and t of f_t and of $\partial f_t / \partial t$, which is not the case of our example. Consequently, Equation (8) leads to the nonsensical expression:

$$\langle \mathcal{O} \rangle_t = \langle \mathcal{O} \rangle_0 + \int_0^t d\tau \int_{-1}^1 \mathcal{O}(e^{-t}x) dx, \tag{132}$$

which diverges linearly with t, rather than converging to $\mathcal{O}(0)$, because $\mathcal{O}(e^{-t}x) \to \mathcal{O}(0)$. This is related to the fact that $\left\langle \Omega^{f_0} \right\rangle_0 \neq 0$, due to the non differentiability of f_0. If, on the other hand, one has $f_0(x) = 3(1 - x^2)/4$ with $x \in [-1, 1]$, one obtains:

$$\Omega^{f_0}(x) = \frac{1 - 3x^2}{1 - x^2}, \quad \text{which leads to} \quad \left\langle \Omega^{f_0} \right\rangle_0 = 0 \tag{133}$$

Then, taking for instance $\mathcal{O}(x) = x^n$, which can be used for analytic functions, one has:

$$\langle \mathcal{O} \rangle_t = \langle \mathcal{O} \rangle_0 + \frac{3}{4} \int_0^t ds\, e^{-ns} \int_{-1}^1 dx\, x^n (1 - 3x^2) = \frac{3}{(n+1)(n+3)} + \frac{3}{(n+1)(n+3)} \left[e^{-nt} - 1 \right] \tag{134}$$

for even n, which correctly converges to $\mathcal{O}(0) = 0$, and 0 for odd n, which equals $\mathcal{O}(0)$ at all times.

The role of Equation (13) is further clarified noticing that for a finite t one can write:

$$\frac{1}{t} \int_0^t \left\langle (\mathcal{O} \circ S^\tau) \cdot \Omega^{f_0} \right\rangle_0 d\tau = \frac{1}{t} \int_0^t d\tau \int_{\mathcal{M}} (\mathcal{O} \circ S^\tau(\mathbf{x})) \cdot \Omega^{f_0}(\mathbf{x}) f_0(\mathbf{x}) d\mathbf{x}$$
$$= \int_{\mathcal{M}} \left[\frac{1}{t} \int_0^t (\mathcal{O} \circ S^\tau(\mathbf{x})) d\tau \right] \Omega^{f_0}(\mathbf{x}) f_0(\mathbf{x}) d\mathbf{x} \tag{135}$$

Then, assuming the $t \to \infty$ limit and the phase integration can be exchanged in Equation (135), the previous equation yields:

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t \left\langle (\mathcal{O} \circ S^\tau) \cdot \Omega^{f_0} \right\rangle_0 d\tau = \overline{\mathcal{O}} \cdot \left\langle \Omega^{f_0} \right\rangle_0$$

where $\overline{\mathcal{O}}$ is the time average of $\mathcal{O}$. This shows that Equation (9) requires $\left\langle \Omega^{f_0} \right\rangle_0 = 0$, i.e., that either Ω^{f_0} vanishes or changes sign in $\mathcal{M}$.

On the other hand, one realizes that this is not sufficient in general. For instance, despite the correct result (134), the blind application of Equation (5) for the case of Equation (121) with density $3(1 - x^2)/4$ in $[-1, 1]$, yields an incorrect evolved density f_t. In this case, we have:

$$\Omega^{f_0}(S^s x) = \Omega^{f_0}(e^{-s}x) = \frac{e^{2s} - 3x^2}{e^{2s} - x^2} \tag{136}$$

and

$$\Omega^{f_0}_{-t,0}(x) = \int_{-t}^0 \Omega^{f_0}(S^s x) ds = 3t + \ln \frac{|e^{-2t} - x^2|}{1 - x^2} \tag{137}$$

which implies:

$$f_t(x) = \frac{3}{4} \left| e^{-2t} - x^2 \right| e^{3t} \tag{138}$$

that is not normalized on $[-1, 1]$. This difficulty is not related to the violation of Equation (13), but to the fact that the support of f_t changes with time or, equivalently, that Equation (5) attributes a positive probability to points outside $[-1, 1]$. An analogous, even more macroscopic, situation arises if $\mathcal{M} = [1, 2]$ and $f_0(x) = 2(3 - x)/3$, which yields $V(1)f_0(1) = V(2)f_0(2)$, hence yields (13) with non vanishing $V(1), V(2)$ and $f_0(1), f_0(2)$. This stresses that dynamical equations and phase space must be properly chosen.

The point is that the preimages $S^{-s}x$ of part of the points in $\mathcal{M}$, for some $s > 0$, lie outside $\mathcal{M}$. In these cases, Equation (5) requires that only the contributions of the points x with $S^{-s}x \in \mathcal{M}$ be included in the calculations, as discussed in Ref. [32].

Example 4. *Consider the dynamical system*

$$\dot{x} = x - x^3, \tag{139}$$

that has two asymptotically stable fixed points $x_- = -1$ and $x_+ = 1$, and an unstable fixed point, $x_0 = 0$. Take the invariant set $\mathcal{M} = [-1, 1]$ as the phase space of the system. The flow is given by:

$$S^t x = \frac{x e^{-t}}{\sqrt{1 + x^2 (e^{-2t} - 1)}}, \quad x \in \mathcal{M}. \tag{140}$$

Assuming $f_0(x) = 1/2$, uniform on $\mathcal{M}$, the dissipation function is expressed by: $\Omega^{f_0}(x) = -\Lambda(x) = 3x^2 - 1$, hence it changes sign in $\mathcal{M}$. The evolved distribution is given by:

$$f_t(x) = \frac{1}{2} \chi_{[-1,1]}(x) e^{-t} (1 - x^2 + x^2 e^{-2t})^{-\frac{3}{2}}. \tag{141}$$

If, on the other hand, $f_0(x) = \frac{3}{4}(1 - x^2)$, we have

$$\Omega^{f_0}(x) = 5x^2 - 1 \tag{142}$$

that also changes sign, and the density at time t becomes:

$$f_t(x) = \frac{3}{4} e^{-t} (1 - x^2) (1 - x^2 + x^2 e^{-2t})^{-\frac{5}{2}}. \tag{143}$$

From (141) and (143) one can see that in both cases $d\mu_t = f_t(x)dx$ converges weakly to $\frac{1}{2}(\delta_{-1} + \delta_1)$. This is correctly so because the flow vanishes at the boundaries of the phase space: $\mathcal{M}$ is invariant; hence the points inside $\mathcal{M}$ have no preimage outside $\mathcal{M}$. In this case, the non-differentiability of the uniform density in $\mathbb{R}$ has no bearing on the theory.

Example 5. *Let us consider the one-dimensional dynamical system:*

$$\dot{x} = -x + \epsilon x^3, \quad 0 < \epsilon \le 1, \tag{144}$$

and denote $V_0(x) = -x$ and $V_p(x) = x^3$. This system has three fixed points: $x_0 = 0$, which is asymptotically stable, and the unstable points $x_- = -1/\sqrt{\epsilon}$ and $x_+ = 1/\sqrt{\epsilon}$. We denote by $\mathcal{M}_\epsilon$ the set $[-1/\sqrt{\epsilon}, 1/\sqrt{\epsilon}] \supset [-1, 1]$, which is invariant, being invariant the intervals $[-1/\sqrt{\epsilon}, 0]$ and $[0, 1/\sqrt{\epsilon}]$. Integrating the differential equation, we obtain

$$S_\epsilon^t x = \frac{x e^{-t}}{\sqrt{1 + \epsilon x^2 (e^{-2t} - 1)}}, \quad for \quad x \in \mathcal{M}_\epsilon. \tag{145}$$

Then, expanding with respect to ϵ,

$$S_\epsilon^t x = S_0^t x - \frac{\epsilon}{2} x^3 e^{-t} (e^{-2t} - 1) + O\left(\epsilon^2 x^5 e^{-t} (e^{-2t} - 1)^2\right), \quad x \in \mathcal{M}_\epsilon, t \geq 0, 0 \leq \epsilon \leq 1, \quad (146)$$

where $S_0^t x = x e^{-t}$. The uniform probability on $[-1,1]$ $f_0(x) = 1/2$, has compact support in $\mathcal{M}_\epsilon \supset [-1,1]$. Substituting $\Omega_{V_0}^{f_0}(x) = 1$ and $\Omega_{V_p}^{f_0}(x) = -3x^2$ in Equation (47) yields:

$$
\begin{aligned}
f_t(x;\epsilon) &= \frac{1}{2} \chi_{[-1,1]}(x) \exp\left[\int_{-t}^0 d\tau\right] \exp\left[-3\epsilon \int_{-t}^0 \frac{x^2 e^{-2\tau}}{1 + \epsilon x^2 (e^{-2\tau} - 1)} d\tau\right] \\
&= \frac{1}{2} \chi_{[-1,1]}(x)\, e^t \, (1 - \epsilon x^2 + \epsilon x^2 e^{2t})^{-\frac{3}{2}}.
\end{aligned}
\quad (147)
$$

This allows us to check that the probability measure $d\mu_t^\epsilon(x) = f_t(x;\epsilon)\, dx$ converges weakly to the delta function at $x = 0$ as $t \to \infty$ The linear approximation for small ϵ can be obtained, in analogy with the Hamiltonian case, cf. (54), by truncating the exponential in (47):

$$\tilde{f}_t(\mathbf{x};\epsilon) = f_0(\mathbf{x}) \exp\left[\int_0^t d\tau\, \Omega_{V_0}^{f_0}(S_\epsilon^{-\tau}\mathbf{x})\right]\left[1 + \epsilon \int_0^t d\tau\, \Omega_{V_p}^{f_0}(S_\epsilon^{-\tau}\mathbf{x})\right] \quad (148)$$

which gives:

$$\tilde{f}_t(x;\epsilon) = \frac{1}{2} \chi_{[-1,1]}(x) e^t \left[1 - \frac{3}{2} \ln\left(1 - \epsilon x^2 + \epsilon x^2 e^{2t}\right)\right]. \quad (149)$$

It can also be observed that the linear approximation at $\epsilon = 0$ of the exact evolved density (147),

$$\hat{f}_t(x;\epsilon) = \frac{1}{2} \chi_{[-1,1]}(x)\, e^t \left[1 - \frac{3}{2}\left(1 - \epsilon x^2 + \epsilon x^2 e^{2t}\right)\right], \quad (150)$$

does not coincide with Equation (149), but it can be obtained from Equation (149) approximating $\ln x$ to first order in x.

Given that $\Omega_{V_0 + \epsilon V_p}^{f_0} = -(1 - 3\epsilon x^2)$, the exact response of ensemble averages is given by:

$$\langle \mathcal{O} \rangle_t = \langle \mathcal{O} \rangle_0 - \frac{1}{2} \int_0^t d\tau \int_{-1}^1 \mathcal{O}(S_\epsilon^\tau x)(1 - 3\epsilon x^2) dx, \quad (151)$$

that, on the basis of the previous remarks, converges to $\mathcal{O}(0)$, as $t \to \infty$. As observed in Section 3, the linear approximation analogous to the GK response is obtained by replacing the exact flow $S_\epsilon^t x$ with the linearized one, $S_0^t x = x\, e^{-t}$, in Equation (151).

Now, let A and B be smooth functions, and f_0 a probability density with compact support K in $\mathcal{M}_\epsilon$, containing $x = 0$ in its interior, for instance $K = [-1,1]$. In order to check T-mixing, see (11), let us write:

$$\lim_{t\to\infty}\left[\int_{-1}^1 A(S_\epsilon^t x)B(x)f_0(x)dx - \int_{-1}^1 A(S_\epsilon^t x)f_0(x)dx \int_{-1}^1 B(x)f_0(x)dx\right] \quad (152)$$

$$= A(0) \int_{-1}^1 B(x)f_0(x)dx - A(0) \int_{-1}^1 B(x)f_0(x)dx = 0 \quad (153)$$

which follows from the uniform convergence of $S_\epsilon^t x$ to 0, from the boundedness of A and B in any compact set and from the continuity of A. This proves the T-mixing property (11) for the present example. The same reasoning shows that the system is also ΩT-mixing.

Example 6. *Consider the one-dimensional dynamical system:*

$$\dot{x} = x - \epsilon x^3, \quad 0 \leq \epsilon \leq 1, \quad (154)$$

that is the time reversed version of (144). *The point* $x_0 = 0$ *is now unstable, while* $x_- = -1/\sqrt{\epsilon}$ *and* $x_+ = 1/\sqrt{\epsilon}$ *are stable, and* $\mathcal{M}_\epsilon = [-1/\sqrt{\epsilon}, 1/\sqrt{\epsilon}] \supset [-1, 1]$ *is invariant. The flow is expressed by:*

$$S_\epsilon^t x = \frac{x e^t}{\sqrt{1 + \epsilon x^2 (e^{2t} - 1)}} = S_0^t x - \frac{\epsilon}{2} x^3 e^t (e^{2t} - 1) + O\left(\epsilon^2 x^5 e^t (e^{2t} - 1)^2\right), \quad \text{for} \quad x \in \mathcal{M}_\epsilon. \quad (155)$$

where $S_0^t x = x e^t$. *The* ϵ *expansion at* $\epsilon = 0$

$$S_\epsilon^t x = S_0^t x - \frac{\epsilon}{2} x^3 e^t (e^{2t} - 1) + \dots \quad (156)$$

has an error that grows without limit as t increases, which shows the unsuitability of the linear response. In this case, always assuming $f_0(x) = 1/2$ *in* $[-1, 1]$, *Equation* (47) *yields*

$$f_t(x; \epsilon) = \frac{1}{2} \chi_{[-1,1]}(x) \exp\left[-\int_{-t}^0 d\tau\right] \exp\left[3\epsilon \int_{-t}^0 \frac{x^2 e^{2\tau}}{1 + \epsilon x^2 (e^{2\tau} - 1)} d\tau\right]$$
$$= \frac{1}{2} \chi_{[-1,1]}(x) e^{-t} (1 - \epsilon x^2 + \epsilon x^2 e^{-2t})^{-\frac{3}{2}}, \quad (157)$$

being $\Omega_{V_0}^{f_0}(x) = -1$ *and* $\Omega_{V_p}^{f_0}(x) = 3x^2$, *since* $V_0(x) = x$ *and* $V_p(x) = -x^3$. *Observe that* $f_t(\pm 1/\sqrt{\epsilon}; \epsilon) \to \infty$ *while* $f_t(x; \epsilon) \to 0$ *for* $-1/\sqrt{\epsilon} < x < 1/\sqrt{\epsilon}$ *as* $t \to \infty$. *Thus the measure* $d\mu_t^\epsilon(x) = f_t(x; \epsilon) \, dx$ *converges weakly to* $\frac{1}{2}\delta_{-1/\sqrt{\epsilon}} + \frac{1}{2}\delta_{1/\sqrt{\epsilon}}$.

Example 7. *Let us consider*

$$\dot{x} = \epsilon(x - x^3), \quad -1 \le \epsilon \le 1, \quad (158)$$

and assume the invariant interval $\mathcal{M} = [-1, 1]$ *as the phase space. For* $\epsilon \ne 0$ *the fixed points are* $x_0 = 0$, $x_- = -1$ *and* $x_+ = 1$. *For* $\epsilon > 0$, *the point* 0 *is unstable, while* -1 *and* 1 *are stable. The opposite holds for* $\epsilon < 0$. *For* $\epsilon = 0$, *all the points are fixed. The flow is given by:*

$$S_\epsilon^t x = \frac{x e^{\epsilon t}}{\sqrt{1 + \epsilon x^2 (e^{2\epsilon t} - 1)}}. \quad (159)$$

Taking the initial uniform distribution on $[-1, 1]$, *being* $\Omega_V^{f_0} = -\epsilon(1 - 3x^2)$, *the evolved density is:*

$$f_t(x; \epsilon) = \frac{1}{2} \chi_{[-1,1]}(x) e^{-\epsilon t} (1 - \epsilon x^2 + \epsilon x^2 e^{-2\epsilon t})^{-\frac{3}{2}}. \quad (160)$$

For $\epsilon > 0$ *the distribution* $d\mu_t^\epsilon(x) = f_t(x; \epsilon) \, dx$ *converges to* $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$, *while for* $\epsilon < 0$ *the limit is* δ_0. *The response is*

$$\langle \mathcal{O} \rangle_t = \langle \mathcal{O} \rangle_0 - \frac{\epsilon}{2} \int_0^t d\tau \int_{-1}^1 \mathcal{O}(S_\epsilon^\tau x)(1 - 3\epsilon x^2) dx \quad (161)$$

Once again, the linear response approach is scarcely interesting.

7. Concluding Remarks

In this paper we have reviewed the developments that led from nonequilbrium molecular dynamics (NEMD) to the fluctuation relations (FR) and to the exact response theory, which develops the previous formalism based on the so-called Transient Time Correlation Function [6]. Within this framework, some authors have adopted the phase space contraction rate $-\Lambda$ as a definition of the steady state entropy production rate σ [23,41], after $-\Lambda$ and σ had been observed to be proportional to each other, in the original paper on fluctuations relations [36]. That paper dealt with the Gaussian thermostatted iso-energetic model of a shearing fluid known as SLLOD. In that case, the steady phase space contraction and entropy production rates could indeed be directly related to each other,

something that in NEMD models had already been pointed earlier [19,51,52]. Mostly the interest had been on steady state state averages of such quantities. It was then realized that such an identification is justified only for steady state averages of certain NEMD models and that, in general, fluctuations require $-\Lambda$ to be distinguished from the energy dissipation rate [24,25,53]. At the same time, the function that represents energy dissipation in molecular dynamics has been identified as the dissipation function Ω, the function that obeys the FR. For the transient FR this happens quite generally; for the steady state FR it is required that correlations w.r.t. the initial distribution do not grow too fast, something implied by the T-mixing condition [26].

Through a number of examples, including NEMD models as well as simple exactly solvable models, we have illustrated these facts and how the exact response formula, based on Ω, applies. In particular we have analyzed the relation between Ω and $-\Lambda$, noting that in the analysis of generic dynamical systems the two quantities are by and large equivalent, while in the case of particle systems the first corresponds to a real observable. We have then compared the exact response formula with the one of the Green-Kubo linear theory, quantifying their difference, that typically grows with the size of the perturbation and of time, but also showing that in some simple case they yield the same result.

We have then investigated "formal thermodynamics" along the lines of Ref. [29], introducing a formal temperature and a formal energy flow, and using Ω instead of $-\Lambda$ as a notion of formal entropy production rate. Analogously to earlier works, e.g., Refs. [54–56], this means that the phase space is treated as real space, and its different regions as different locations in real space. Clearly, the $2dN$-dimensional phase space of an N particle system in d dimensions and the phase points that flow in it are totally different from the $2d$-dimensional real space and the particles that flow through it [43,44]. (A point in phase space has no extension nor "inertia", it does not interact with other points, and it equally represents the microstate of a large or small system. Any other point represents another system that can be arbitrarily close or far from it. A particle in real space has a finite extension, interacts with nearby particles, and follows Newton's laws with its inertia. Just this has numerous consequences. For instance, while the Gibbs entropy represents thermodynamic entropy only at equilibrium, the Boltzmann entropy works also in time dependent situations, Refs. [44,57]). Therefore, in general, a direct identification of flows and related quantities in phase space with their physical counterparts is not possible. Nevertheless, this analysis is reminiscent of the thermodynamic formalism, which is important for dynamical systems in general [33]. Furthermore, it may find direct physical application in systems of non-interacting particles, for which the factorization of the phase space distributions makes phase space and real space almost the same. Such systems may not behave thermodynamically, but they are of greater and greater interest in e.g., bio- and nano-technology (cf. transport in highly confining media), for photons in optical wave-guides or random media, for cold atoms etc.

Finally, we have observed that the exact response formula (8), mainly rests on the regularity of the vector fields, and on the ergodic consistency of the initial distribution, which ensure the regularity of the dissipation function as well. These mathematical conditions need further scrutiny, in more realistic cases, but are relatively minimal, hence hint at a wide applicability of the theory. The investigation of physical applications of this theory is just at the beginning; many are indeed the cases in which perturbations cannot be taken as small compared to the reference signals; they range from nano-technology to geophysical and even astrophysical phenomena [58].

References

1. Fermi, E.; Pasta, P.; Ulam, S.; Tsingou, M. *Studies of the Nonlinear Problems*; Technical report; Los Alamos Scientific Laboratory: Los Alamos, NM, USA, 1955.
2. Gallavotti, G. *The Fermi-Pasta-Ulam Problem: A Status Report*; Springer: New York, NY, USA, 2007; Volume 728.
3. Allen, M.P.; Tildesley, D.J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, UK, 2017.
4. Battimelli, G.; Ciccotti, G.; Greco, P. *Il Computer Incontra la Fisica Teorica*; Carocci Editore: Rome, Italy, 2020.
5. Evans, D.J.; Searles, D.J.; Williams, S.R. *Fundamentals of Classical Statistical Thermodynamics*; Wiley Online Library: Hoboken, NJ, USA, 2016.
6. Evans, D.J.; Morriss, G.P. *Statistical Mechanics of Nonequilbrium Liquids*; Cambridge University Press: Cambridge, UK, 2008.
7. Todd, B.D.; Daivis, P.J. *Nonequilibrium Molecular Dynamics: Theory, Algorithms and Applications*; Cambridge University Press: Cambridge, UK, 2017.
8. Evans, D.J.; Hoover, W.G.; Failor, B.H.; Moran, B.; Ladd, A.J.C. Nonequilibrium molecular dynamics via gauss's principle of least constraint. *Phys. Rev. A* **1983**, *28*, 1016. [CrossRef]
9. Jepps, O.G.; Rondoni, L. Deterministic thermostats, theories of nonequilibrium systems and parallels with the ergodic condition. *J. Phys. A* **2010**, *43*, 133001. [CrossRef]
10. Plimpton, S.; Crozier, P.; Thompson, A. *Lammps User's Manual*; Sandia National Laboratory: Albuquerque, NM, USA; Livermore, CA, USA, 2005.
11. Van der Spoel, D.; van Buuren, A.R.; Apol, E.; Meulenhoff, P.J.; Tieleman, D.P.; Sijbers, A.L.T.M.; Feenstra, K. *The Gromacs Manual (Version 3.3)*; Groningen University: Groningen, The Netherlands, 2004.
12. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H.J.C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341. [CrossRef]
13. Chibbaro, S.; Rondoni, L.; Angelo, V. *Reductionism, Emergence And Levels Of Reality*; Springer: New York, NY, USA, 2016.
14. Evans, D.J. The equivalenceof norton and th venin ensembles. *Mol. Phys.* **1993**, *80*, 221. [CrossRef]
15. Gallavotti, G. *Nonequilibrium and Irreversibility*; Springer: New York, NY, USA, 2014.
16. Rondoni, L. Deterministic thermostats and flctuation relations. In *Dynamics of Dissipation*; Springer: New York, NY, USA, 2002; pp. 35–61.
17. Saint-Raymond, L. From the BGK model to the Navier-Stokes equations. *Ann. Sci. l'Ecole Normale Supérieure* **2003**, *36*, 271–317.
18. Villani, C. A review of mathematical topics in collisional kinetic theory. *Handb. Math. Fluid Dyn.* **2002**, *1*, 3–8.
19. Evans, D. Response theory as a free-energy extremum. *Phys. Rev. A* **1985**, *32*, 2923. [CrossRef]
20. Andrey, L. The rate of entropy change in non-hamiltonian systems. *Phys. Lett. A* **1985**, *111*, 45–46. [CrossRef]
21. Searles, D.J.; Evans, D.J. Fluctuation theorem for stochastic systems. *Phys. Rev. E* **1999**, *60*, 159. [CrossRef]
22. Gallavotti, G.; Cohen, E.G.D. Dynamical ensembles in nonequilibrium statistical mechanics. *Phys. Rev. Lett.* **1995**, *74*, 2694. [CrossRef]
23. Ruelle, D. Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics. *J. Stat. Phys.* **1999**, *95*, 393–468. [CrossRef]
24. Cohen, E.G.D.; Rondoni, L. Note on phase space contraction and entropy production in thermostatted hamiltonian systems. *CHAOS* **1998**, *8*, 357. [CrossRef] [PubMed]
25. Searles, D.J.; Evans, D.J. Ensemble dependence of the transient fluctuation theorem. *J. Chem. Phys.* **2000**, *113*, 3503–3509. [CrossRef]
26. Evans, D.J.; Williams, S.R.; Searles, D.J.; Rondoni, L. On typicality in nonequilibrium steady states. *J. Stat. Phys.* **2016**, *164*, 842–857. [CrossRef]
27. Jepps, O.G.; Rondoni, L. A dynamical-systems interpretation of the dissipation function, t-mixing and their relation to thermodynamic relaxation. *J. Phys. A* **2016**, *49*, 154002. [CrossRef]
28. Evans, D.J.; Morriss, G.P. Statistical mechanics: A short treatise. *Phys. Today* **2001**, *54*, 63.
29. Qian, H.; Wang, S.; Yi, Y. Entropy productions in dissipative systems. *Proc. Am. Math. Soc.* **2019**, *147*, 5209–5225. [CrossRef]
30. Searles, D.J.; Rondoni, L.; Evans, D.J. The steady state fluctuation relation for the dissipation function. *J. Stat. Phys.* **2007**, *128*, 1337–1363. [CrossRef]

31. Jaksic, V.; Pillet, C.-A.; Rey-Bellet, L. Entropic fluctuations in statistical mechanics: I. classical dynamical systems. *Nonlinearity* **2011**, *24*, 699.

32. Rondoni, L.; Dematteis, G. Physical ergodicity and exact response relations for low-dimensional maps. *Comput. Methods Sci. Technol.* **2016**, *22*, 71. [CrossRef]

33. Ruelle, D. *Thermodynamic Formalism*; Cambridge University Press: Cambridge, UK, 2004.

34. Morawetz, K. Nonequilibrium thermodynamics with binary quantum correlations. *Phys. Rev. E* **2017**, *96*, 032106. [CrossRef]

35. Morawetz, K. *Interacting Systems Far From Equilibrium: Quantum Kinetic Theory*; Oxford University Press: Oxford, UK, 2018.

36. Evans, D.J.; Cohen, E.G.D.; Morriss, G.P. Probability of second law violations in shearing steady states. *Phys. Rev. Lett.* **1993**, *71*, 2401. [CrossRef] [PubMed]

37. Bonetto, F.; Gallavotti, G.; Giuliani, A.; Zamponi, F. Chaotic hypothesis, fluctuation theorem and singularities. *J. Stat. Phys.* **2006**, *123*, 39. [CrossRef]

38. Evans, D.J.; Searles, D.J.; Rondoni, L. Application of the gallavotti-cohen fluctuation relation to thermostated steady states near equilibrium. *Phys. Rev. E* **2005**, *71*, 056120. [CrossRef] [PubMed]

39. Vilela-Mendes, R.; Duarte, J. Decomposition of vector fields and mixed dynamics. *J. Math. Phys.* **1981**, *22*, 1420. [CrossRef]

40. Rondoni, L. Introduction to nonequilibrium statistical physics and its foundations. In *Lecture Notes at 2019 International Graduate Summer School on Soft Matter and Nonequilibrium Physics*; Xiamen University and Huaqiao University: Xiamen, China, 2020.

41. Gallavotti, G.; Cohen, E.G.D. Dynamical ensembles in stationary states. *J. Stat. Phys.* **1995**, *80*, 931. [CrossRef]

42. Falcioni, M.; Palatella, L.M.; Pigolotti, S.; Rondoni, L.; Vulpiani, A. Ergodic properties of high-dimensional symplectic maps. *Phys. Rev. A* **1991**, *44*, 2263. [CrossRef]

43. Rondoni, L.; Cohen, E.G.D. On some derivations of irreversible thermodynamics from dynamical systems theory. *Phys. D Nonlinear Phenom.* **2002**, *168*, 341–355. [CrossRef]

44. Rondoni, L.; Cohen, E.G.D. Gibbs entropy and irreversible thermodynamics. *Nonlinearity* **2000**, *13*, 1905. [CrossRef]

45. Marconi, U.M.B.; Puglisi, A.; Rondoni, L.; Vulpiani, A. Fluctuation-dissipation: Response theory in statistical physics. *Phys. Rep.* **2008**, *461*, 111. [CrossRef]

46. Tuckerman, M.E. *Statistical Mechanics: Theory and Molecular Simulation*; Oxford University Press: Oxford, UK, 2010.

47. Kubo, R.; Toda, M.; Hashitsume, N. *Statistical Physics II*; Springer: New York, NY, USA, 1985.

48. van Kampen, N. The case against linear response theory. *Phys. Nor.* **1971**, *5*, 279.

49. Bianucci, M.; Mannella, M.; Grigolini, P. Linear response of hamiltonian chaotic systems as a function of the number of degrees of freedom. *Phys. Rev. Lett.* **1996**, *77*, 1258. [CrossRef] [PubMed]

50. Colangeli, M.; Rondoni, L.; Vulpiani, A. Fluctuation-dissipation relation for chaotic non-hamiltonian systems. *JSTAT* **2012**, *5*, L04002. [CrossRef]

51. Evans, D.J.; Cohen, E.G.D.; Morriss, G.P. Viscosity of a simple fluid from its maximal lyapunov exponents. *Phys. Rev. A* **1990**, *42*, 5990. [CrossRef] [PubMed]

52. Hoover, W.; Posch, H. Direct measurement of equilibrium and nonequilibrium lyapunov spectra. *Phys. Lett. A* **1987**, *123*, 227. [CrossRef]

53. Klages, R. *Microscopic Chaos, Fractals and Transport in Nonequilibrium Statistical Mechanics*; World Scientific: Singapore, 2007.

54. Gaspard, P. Entropy production in open volume-preserving systems. *J. Stat. Phys.* **1997**, *88*, 1215. [CrossRef]

55. Gaspard, P. *Chaos, Scattering and Statistical Mechanics*; Cambridge University Press: Cambridge, UK, 1998.

56. Gilbert, T.; Dorfman, J.R. Entropy production: From open volume preserving to dissipative systems. *J. Stat. Phys.* **1999**, *96*, 225. [CrossRef]

57. Rondoni, L.; Pigolotti, S. On $\Gamma-$ and $\mu-$space descriptions: Gibbs and boltzmann entropies of symplectic coupled maps. *Phys. Scr.* **2012**, *86*, 1. [CrossRef]
58. Baiesi, M.; Burigana, C.; Conti, L.; Falasco, G.; Maes, C.; Rondoni, L.; Trombetti, T. Possible nonequilibrium imprint in the cosmic background at low frequencies. *Phys. Rev. Res.* **2020**, *2*, 013210. [CrossRef]