Physics Reports 480 (2009) 57–159

Contents lists available at ScienceDirect

Physics Reports

journal homepage: www.elsevier.com/locate/physrep

Statistical mechanics and dynamics of solvable models with long-range interactions

Alessandro Campa [a,*], Thierry Dauxois [b], Stefano Ruffo [c]

[a] Complex Systems and Theoretical Physics Unit, Health and Technology Department, Istituto Superiore di Sanità, and INFN Roma 1, Gruppo Collegato Sanità, Viale Regina Elena 299, 00161 Roma, Italy

[b] Université de Lyon, Laboratoire de Physique de l'École Normale Supérieure de Lyon, CNRS, 46 allée d'Italie, 69364 Lyon cedex 07, France

[c] Dipartimento di Energetica "S. Stecco" and CSDC, Università di Firenze, INFN, Via S. Marta, 3 I-50139, Firenze, Italy

ARTICLE INFO

Article history:
Accepted 9 July 2009
Available online 22 July 2009
editor: H. Orland

PACS:
05.20.-y
05.20.Dd
05.20.Gg
64.60.Bd
64.60.De

Keywords:
Long-range interactions
Ensemble inequivalence
Negative specific heat
Ergodicity breaking
Vlasov equation
Quasi-stationary states

ABSTRACT

For systems with long-range interactions, the two-body potential decays at large distances as $V(r) \sim 1/r^{\alpha}$, with $\alpha \leq d$, where d is the space dimension. Examples are: gravitational systems, two-dimensional hydrodynamics, two-dimensional elasticity, charged and dipolar systems. Although such systems can be made extensive, they are intrinsically *non additive*: the sum of the energies of macroscopic subsystems is not equal to the energy of the whole system. Moreover, the space of accessible macroscopic thermodynamic parameters might be *non convex*. The violation of these two basic properties of the thermodynamics of short-range systems is at the origin of *ensemble inequivalence*. In turn, this inequivalence implies that specific heat can be negative in the microcanonical ensemble, and temperature jumps can appear at microcanonical first order phase transitions. The lack of convexity allows us to easily spot regions of parameter space where ergodicity may be broken. Historically, negative specific heat had been found for gravitational systems and was thought to be a specific property of a system for which the existence of standard equilibrium statistical mechanics itself was doubted. Realizing that such properties may be present for a wider class of systems has renewed the interest in long-range interactions. Here, we present a comprehensive review of the recent advances on the statistical mechanics and out-of-equilibrium dynamics of solvable systems with long-range interactions. The core of the review consists in the detailed presentation of the concept of ensemble inequivalence, as exemplified by the exact solution, in the microcanonical and canonical ensembles, of mean-field type models. Remarkably, the entropy of all these models can be obtained using the method of large deviations. Long-range interacting systems display an extremely slow relaxation towards thermodynamic equilibrium and, what is more striking, the convergence towards *quasi-stationary states*. The understanding of such unusual relaxation process is obtained by the introduction of an appropriate kinetic theory based on the Vlasov equation. A statistical approach, founded on a variational principle introduced by Lynden-Bell, is shown to explain qualitatively and quantitatively some features of quasi-stationary states. Generalizations to models with both short and long-range interactions, and to models with weakly decaying interactions, show the robustness of the effects obtained for mean-field models.

* Corresponding author.
E-mail addresses: campa@iss.infn.it (A. Campa), Thierry.Dauxois@ens-lyon.fr (T. Dauxois), Stefano.Ruffo@unifi.it (S. Ruffo).

0370-1573/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.physrep.2009.07.001

Contents

1. Introduction

A wide range of problems in physics concerns systems with long-range interactions. However, their statistical and dynamical properties are much less understood than those of short-range systems [1,2]. One finds examples of long-range interacting systems in astrophysics [3,4], plasma physics [5], hydrodynamics [6,7], atomic physics [8], and nuclear physics [9]. This ubiquitous presence in different physics disciplines alone would itself justify the need for a general and interdisciplinary understanding of the physical and mathematical problems raised by long-range interacting systems.

In this review, we are interested in systems with a large number N of degrees of freedom, for which a statistical physics approach is mandatory, independently of the specific features of the interactions. Therefore, we will discuss in the following equilibrium and out-of-equilibrium properties of long-range systems relying on the tools of statistical mechanics [10].

Let us define which is the property of the interaction that makes it short or long-ranged. We consider systems where the interaction potential is given by the sum, over pairs of the elementary constituents, of a two-body translationally invariant potential. For sufficiently large distances r, the absolute value of the two-body potential is bounded by $r^{-\alpha}$. If the positive power α is larger than the dimension d of the space where the system is embedded, $\alpha > d$, we define the system to be short-range. Otherwise, if $\alpha \leq d$, the system is long-range. The reason for this definition is that, in the large N limit all the mathematical and physical differences between short and long-range systems can be traced back to this property of the interaction potential. We should remark that this definition of the range of the interaction does not coincide with others, where the range is instead defined by a characteristic length appearing in the interaction potential. In this latter definition, any interaction decaying as a power law at large distances, thus without characteristic length, is considered as long range. However, for the physical and mathematical problems found in the statistical mechanics of many-body systems, the definition used throughout this review is more appropriate, since, as we have stressed, it is related to the interaction property that determines the behavior of such systems for large N.

Our purpose will be to illustrate, especially through the use of simple models, the peculiar properties of long-range systems, and the tools and techniques that are employed to describe them. To give a flavor of the issues that will be considered, we would like to begin with a very simple description of the main problems that one faces in the study of these systems.

The aim of equilibrium statistical mechanics is to derive the thermodynamic properties of a macroscopic system from microscopic interactions [10]. The connection between micro and macro is realized through the introduction of statistical ensembles. Different thermodynamic potentials describe situations in which different thermodynamic parameters are used in the characterization of the system, and, in the aforementioned connection with the microscopic interactions, different statistical ensembles are related to different thermodynamic potentials. However, it is usually stated (and experimentally verified for many physical systems) that, as far as macroscopic averages are concerned, i.e. in the thermodynamic limit ($N \to \infty$, $V \to \infty$ with $N/V = const.$), the predictions of statistical mechanics do not depend on the chosen ensemble. *Ensemble equivalence* is related to the fact that, given a sufficient number of macroscopic thermodynamic parameters (two in a one-component system), the others are fixed in the large volume limit, apart from vanishingly small relative fluctuations.

An important feature of long-range systems is that *ensembles can be inequivalent* [4,11–14], and therefore one of the main issues in the statistical mechanics of these systems is a careful examination of the relations between the different ensembles, in particular of the conditions that determine their equivalence or inequivalence. We emphasize that ensemble inequivalence is not merely a mathematical drawback, but it is the cause, as it will be shown in this review, of physical properties of these systems that can be experimentally verified. Probably, one of the most striking features of long-range systems is the possibility to display *negative specific heat* in the microcanonical ensemble [11,12,15–18]. Specific heat is always positive in the canonical ensemble, independently of the nature of the interactions, since it is given by the expectation value of a positive quantity. It turns out that microcanonical equilibrim contains all the information about canonical equilibrium, while the converse is wrong in case of ensemble inequivalence [19,20]. The discrepancy between the two ensembles extends to other observables related to the response of the system to a change in a thermodynamic parameter: a concrete example will be given discussing magnetic susceptibility.

As far as out-of-equilibrium dynamical properties are concerned, many-body long-range systems again show peculiar behaviors. The approach to equilibrium of short-range systems is usually characterized by the time scales that govern the equations of motion of the elementary constituents [21]. For systems without disorder, these time scales are typically small when compared to the observational time scales. Sometimes these systems can be trapped in metastable states that last for a long time. These states are local extrema of thermodynamic potentials and, in practical cases, their realization requires a very careful preparation of the system (e.g. undercooled liquids and superheated solids). If perturbed, the system

rapidly converges towards the equilibrium state. It can also happen that the relaxation time depends on the volume if hydrodynamics modes are present or in coarsening processes.

For long-range systems, dynamics can be extremely slow and the approach to equilibrium can take a very long time, that increases with the number N of elementary constituents [22]. This feature is induced by the long-range nature of the interaction itself and is not a consequence of the existence of a collective phenomenon. The state of the system during this long transient is *quasi-stationary* [23–26], since its very slow time evolution allows us to define slowly varying macroscopic observables, like for local equilibrium or quasi-static transformations. It should be however remarked that quasi-stationary states are not thermodynamic metastable states, since they do not lie on local extrema of equilibrium thermodynamic potentials. The explanation of their widespread presence should rely only upon the dynamical properties of the systems. It must be stressed that the nature of quasi-stationary states can be strongly dependent on the initial condition, as it will become clear from the examples that we will give. In addition, a variety of macroscopic structures can form spontaneously in out-of-equilibrium conditions for isolated systems: a fact that should not be a surprise given that already the equilibrium states of long-range systems are usually inhomogeneous. In short-range systems, macroscopic structures can arise as an effect of an external forcing (Rayleigh–Bénard convection, Benjamin–Feir instability, Faraday waves) [27] or due to the nonlinearity of the governing equations of motion (solitons, breathers) [28], but are usually strongly selected by the specific dynamical properties and by the geometrical conditions. All this shows the great richness of the dynamics of long-range systems.

Summarizing, a satisfying theoretical framework concerning the behavior of system with long-range interactions should necessarily address the following aspects:

- For what concerns equilibrium properties, the determination of the physical conditions that determine equivalence or inequivalence of the statistical ensembles and, for the latter case, the relation between macrostates in the different statistical ensembles.
- As for the out-of-equilibrium features, the development of a consistent kinetic theory able to explain the formation of quasi-stationary states and their final relaxation to equilibrium.

Although different aspects of systems with long-range interactions have been studied in the past in specific scientific communities, notably astrophysics and plasma physics, this has not constituted a seed for more general theoretical studies. In the last decade or so, it has become progressively clearer that the ubiquitous presence of long-range forces needs an approach that integrates different methodologies [1,2]. This has induced a widespread interest in long-range systems throughout numerous research groups. The successive development has led to a better understanding of both the equilibrium and out-of-equilibrium properties of such systems. The time is ripe to summarize what is known on firm basis about the equilibrium statistical mechanics of systems with long-range interactions and to describe the preliminaries of a theory of non-equilibrium.

In this review, we have chosen to present the main problems, tools and solutions by discussing paradigmatic examples, which are simple and general enough to be useful also for specific applications. Therefore, we will mostly emphasize:

- the equilibrium statistical mechanics solutions of simple mean-field toy models, which is a first step for understanding phase diagrams of more complex long-range systems;
- the basic ingredients of a kinetic theory of a model able to catch the essential properties of out-of-equilibrium dynamics.
- the analysis of the phase diagram of models with both short and long-range interactions, or with forces that weakly decay in space.

The structure of the review is the following. In Section 2, we introduce the subject presenting the definition of long-range interactions and discussing the non-additivity property and its consequences. In Section 3, we briefly review the relevant physical systems with long-range interactions: gravitational systems, 2D hydrodynamics, 2D elasticity, charged systems, dipolar systems and small systems. Section 4 constitutes the core of the review. We present there the equilibrium properties of several mean-field models for which one can analytically compute both the free energy and the entropy. The solution is obtained by different methods, including the powerful large deviation method, which allows one to solve models with continuous variables. The dynamics is tackled in Section 5, where the main result reviewed is the widespread presence of quasi-stationary states that hinder the relaxation to a Boltzmann–Gibbs equilibrium. This phenomenon is strongly supported by numerical simulations and can be studied using kinetic equations explicitly devised for long-range interactions (Vlasov, Lenard–Balescu). Few analytical results exist for non-mean-field models: we collect some of them in Section 6. The message is that the introduction of short-range terms does not spoil the features of mean-field models and that some weakly decaying interactions can be treated rigorously. Finally, we draw conclusions and we discuss some perspectives in Section 7.

2. The additivity property and the definition of long-range systems

2.1. Extensivity vs. additivity

In order to easily illustrate the issues of *extensivity* and *additivity*, it is useful to consider a concrete example and restrict ourselves to the energy as the thermodynamic extensive variable. We employ a very simple model that is used in the study

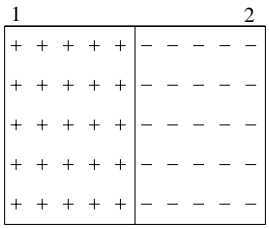


Fig. 1. Schematic picture of a system separated in two equal parts with $N/2$ spins up in domain 1 and $N/2$ spins down in domain 2.

of magnetic systems, namely the Curie–Weiss Hamiltonian

$$H_{CW} = -\frac{J}{2N}\sum_{i,j=1}^{N} S_i S_j = -\frac{J}{2N}\left(\sum_{i=1}^{N} S_i\right)^2, \tag{1}$$

where the spins $S_i = \pm 1$ are attached to sites labeled by $i = 1, \ldots, N$. In this example, the interaction does not decay at all with the distance: indeed, each spin interacts with equal strength with all the other spins. Such systems are usually referred to as mean-field systems. With the $1/N$ prefactor in (1), the total energy increases as N, then the energy per spin converges to a finite value in the thermodynamic limit, which is a physically reasonable requirement. We recall that one can find rigorous definitions of the thermodynamic limit in Ruelle's book [29]. Model (1) is *extensive*: for a given intensive magnetization

$$m = \frac{\sum_i S_i}{N} = \frac{M}{N}, \tag{2}$$

where M is the extensive magnetization, if one doubles the number of spins the energy doubles. On the other hand, Hamiltonian (1) is not additive, in spite of the presence of the regularizing factor $1/N$. Indeed, let us divide the system, schematically pictured in Fig. 1, in two equal parts. In addition, let us consider the particular case in which all spins in the left part are equal to $+1$, whereas all spins in the right part are equal to -1. The energy of the two parts is $E_1 = E_2 = -JN/4$. However, if one computes the total energy of the system, one gets $E = 0$. Since $E \neq E_1 + E_2$, such a system is not additive, at least for this configuration. One could easily generalize the argument to generic configurations. The problem is not solved by increasing system size N, because the "interaction" energy $E_{1,2} = E - E_1 - E_2 = JN/2$ increases with N.

Extensivity in the Curie–Weiss model is provided by the $1/N$ prefactor in Hamiltonian (1), which makes the energy proportional to N. This is the so-called Kac prescription [30]. This energy rescaling guarantees a competition between energy E and entropy S, which is crucial for phase transitions. Indeed, introducing the free energy $F = E - TS$, where T is temperature, the Kac prescription implies that the two competing terms on the r.h.s. both scale as N, since temperature is intensive and entropy scales as N. This latter scaling deserves a further analysis which will be developed later on. Intuitively, the fact that the interaction is long-range does not alter the density of states, which usually grows factorially with N. An alternative prescription would be to make the temperature extensive $T \to TN$, in such a way that the energy and entropy terms of the free energy both scale as N^2. The two prescriptions give equivalent physical consequences. In systems with kinetic energy, being temperature the average kinetic energy per particle, rescaling temperature corresponds to a renormalization of velocities, and finally of the time scale.

After having defined the two distinct concepts of *extensivity* and *additivity* with reference to a specific model, let us clarify more generally these two notions.

Indeed, it is very important, as a first step, to give the general definitions of *extensivity* and *additivity*, and to clarify the distinction between these two concepts. It is convenient to first consider the situation that one encounters in short-range systems. We can imagine to divide a system at equilibrium in two parts occupying equal volumes. Some thermodynamic variables of each half of the system will be equal to the corresponding ones of the total system, others will be halved. Temperature and pressure are examples of the first kind of thermodynamic variable: they do not depend on the size of the system and are called intensive variables. Energy, entropy and free energy are variables of the second kind; their value is proportional to the system size, i.e. to the number of elementary constituents (for given values of the intensive variables), and they are called extensive variables. The property where the size dependent thermodynamic variables are proportional to system size is called *extensivity*, and systems with this property are called extensive. The specific value of extensive variables (e.g., the energy per unit particle, or per unit mass, or per unit volume) gives rise to new intensive quantities. Considering the energy of a system, we see that it has also the property of *additivity*, that consists of the following. Dividing the systems in two macroscopic parts, the total energy E will be equal to $E_1 + E_2 + E_{int}$, with E_i the energy of the i-th part, and E_{int} the interaction energy between the two parts. In the thermodynamic limit the ratio $E_{int}/(E_1 + E_2)$ tends to zero; therefore in this limit $E \approx E_1 + E_2$. This property is called *additivity*; systems with this property (for the energy as well as for other size dependent quantities) are called additive. It is evident that extensivity and additivity are related. Indeed, in the definition of extensivity just given, we could not have concluded that the energy of each part of the system is half the total energy, if the

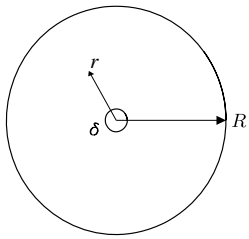


Fig. 2. Schematic picture of the domain considered for the evaluation of the energy ε of a particle. It is a spherical shell of outer radius R and inner radius δ.

interaction energy E_{int} would not be negligible. Additivity implies extensivity (thus non extensivity implies non-additivity), but not the reverse, since the interaction energy might scale with N, as in the Curie–Weiss model. This comment applies more generally to all long-range systems. The unusual properties of these systems derive from the lack of additivity.

2.2. Definition of long-range systems

We have just shown that mean-field systems like the Curie–Weiss model can be made extensive using Kac's trick. However, extensive systems could be non-additive. Here, we will discuss the case of interactions that decay as a power law at large distances. We will show that the energy ε of a particle (excluding self-energy) diverges if the potential does not decay sufficiently fast, implying that the total energy grows superlinearly with volume at constant density, which violates extensivity. These interactions are called *long-range* or *non-integrable*, just referring to this divergence of the energy. Energy convergence can be restored by an appropriate generalization of Kac's trick.

Let us estimate the energy ε by considering a given particle placed at the center of a sphere of radius R where the other particles are homogeneously distributed. We will exclude the contribution to ε coming from the particles located in a small neighborhood of radius δ (see Fig. 2). This is motivated by the necessity to regularize the divergence of the potential at small distances, which has nothing to do with its long-range nature.

If the other particles interact with the given one via a potential that at large distances decays like $1/r^{\alpha}$, we obtain in d-dimensions

$$\varepsilon = \int_{\delta}^{R} \mathrm{d}^d r \rho \frac{J}{r^{\alpha}} = \rho J \Omega_d \int_{\delta}^{R} r^{d-1-\alpha} \mathrm{d}r = \frac{\rho J \Omega_d}{d-\alpha} \left[R^{d-\alpha} - \delta^{d-\alpha} \right], \quad \text{if } \alpha \neq d, \tag{3}$$

where ρ is the generic density (e.g. mass, charge), J is the coupling constant and Ω_d is the angular volume in dimension d (2π in $d = 2$, 4π in $d = 3$, etc.). When increasing the radius R, the energy ε remains finite if and only if $\alpha > d$. This implies that the total energy E will increase linearly with the volume V, i.e. the system is extensive. Such interactions are the usual short-range ones. On the contrary, if $\alpha \leq d$ the energy ε grows with volume as $V^{1-\alpha/d}$ (logarithmically in the marginal case $\alpha = d$). This implies that the total energy E will increase superlinearly, $E \propto V^{2-\alpha/d}$, with volume. However, analogously to the Kac's prescription, one can redefine the coupling constant $J \to JV^{\alpha/d-1}$ and get a perfectly extensive system. Mean-field models, like Hamiltonian (1), correspond to the value $\alpha = 0$, since the interaction does not depend on the distance, and one recovers the usual Kac's rescaling. Cases where the energy grows superlinearly define the long-range nature of the interaction. However, as we have shown for mean-field systems, the fact that energy can be made extensive, does not imply that the system is additive. Implications of the lack of additivity for long-range systems will be discussed throughout the paper.

2.3. Convexity in thermodynamic parameters

Let us consider the space of the extensive thermodynamic parameters. To be specific, as in Fig. 3, we consider the (E, M) plane of energy and magnetization (2). The attainable region in this space is always convex when only short-range interactions are present. This property is a direct consequence of additivity. Consider two different subsystems with two different energies E_1 and E_2, and two different magnetization values M_1 and M_2. Introducing a parameter λ taking values between 0 and 1, depending on the relative size of the subsystems, the system obtained by combining the two subsystems has an energy $E = \lambda E_1 + (1-\lambda)E_2$, and a magnetization $M = \lambda M_1 + (1-\lambda)M_2$. Any value of λ between 0 and 1 is realized thermodynamically, just varying the relative size of the two subsystems. This is exactly the *convexity property* of the attainable region of the space of thermodynamic parameters. In particular, one can consider two subsystems with the same energy but different magnetization. Varying the relative size of the two subsystems, the combined system will have the same energy and any possible magnetization between the two values of the two subsystems. It is important to stress that the convexity property is possible if additivity is satisfied, since the interaction energy between the two subsystems has been neglected. As already remarked, the additivity property is generically valid for large enough and short-range interacting systems. Moreover convexity implies that the space of thermodynamic parameters is connected.

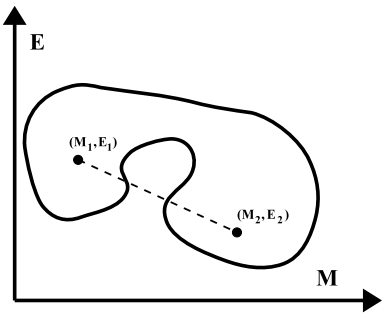


Fig. 3. The set of accessible macrostates in the (M, E) space can have a non-convex shape for systems with long-range interactions, such that if (M_1, E_1) and (M_2, E_2) can be realized macroscopically, this is not necessarily true for all the states joining these two along the straight dashed line.

On the contrary, systems with long-range interactions are not additive, and thus intermediate values of extensive parameters are not necessarily accessible (see Fig. 3). This feature has profound consequences on the dynamics of systems with long-range interactions. Gaps may open up in the space of extensive variables. Since the space of thermodynamic parameters is no longer connected, *ergodicity breaking* might appear when considering continuous microcanonical dynamics of such a system. We will discuss again this question in detail in Section 4.5.4 by emphasizing simple examples.

2.4. Lattice systems

When defining the long-range nature of the interaction, care must be taken of the specific nature of the microscopic variables. These could be divided in two classes: the coordinates related to the translational degrees of freedom (e.g. cartesian coordinates), and those giving the internal state of each particle (e.g. spin variables). Both could be either continuous or discrete. When coordinates take fixed discrete values, one speaks of *lattice systems*. As for the internal degrees of freedom the variables could be discrete (e.g. spin systems) or continuous (e.g. systems of rotators).

For systems with continuous translational degrees of freedom that do not possess internal degrees of freedom, the potential energy can be written in the general form

$$U(\overrightarrow{r}_1, \ldots, \overrightarrow{r}_N) = \sum_{1\le i<j\le N} V(|\overrightarrow{r}_i - \overrightarrow{r}_j|), \tag{4}$$

where $(\overrightarrow{r}_1, \ldots, \overrightarrow{r}_N)$ are cartesian coordinates in d-dimensional space and we assume that the translationally invariant pair potential V depends only on the modulus of the distance between two particles. Systems of gravitational point masses or Coulomb point charges fall into this category.

In a lattice system, for which each site i of a d-dimensional lattice, located by the position vector $\mathbf{r}_i$, hosts a particle possessing one or more internal degrees of freedom, collectively denoted by the vector $\mathbf{q}_i$ (the dimensionality of this vector is independent of d), the potential energy can be written as

$$U(\mathbf{q}_1, \ldots, \mathbf{q}_N) = \sum_{1\le i<j\le N} C_{ij} V(\mathbf{q}_i, \mathbf{q}_j) + g \sum_{i=1}^{N} V_e(\mathbf{q}_i), \tag{5}$$

where the coupling constants C_{ij} are translationally and rotationally invariant (i.e., they depend only on $|\overrightarrow{r}_i - \overrightarrow{r}_j|$). We also allow the presence of an external field that couples to the particles via the function $V_e(\mathbf{q})$ with a strength g. As pointed out above, the variables $\mathbf{q}_i$ may take continuous or discrete values.

We have already discussed in the previous Subsection that long-range systems of type (4) are those for which, at large distance, $V(r) \sim r^{-\alpha}$ with $\alpha \le d$. Similarly, long-range lattice systems (5) can be characterized by a slow decay of the coupling constants C_{ij}. If these latter behave at large distance like $|\overrightarrow{r}_i - \overrightarrow{r}_j|^{-\alpha}$ with $\alpha \le d$, the system is long-range. The energy grows also in this case superlinearly with the volume and one will need to introduce a Kac rescaling factor. At variance with systems of type (4) we need not worry about the behavior at short distances, because the lattice regularizes any possible divergence.

A kinetic energy term can be added to the potential energy one for systems of type (4) and for lattice systems (5) when the variables are continuous.

In this review, we will mostly concentrate our attention on lattice systems. Since they cannot display any short distance singularity, their thermodynamic and dynamical behavior highlights the essential features of long-range interactions. Therefore, these systems are more suitable for presenting an overview of the main results and of the tools used to deal with long-range interactions.

2.5. Non-additivity and the canonical ensemble

Let us recall that, in $d = 3$, neglecting the dependence on the number of particles N and on the volume V, the microcanonical partition function (proportional to the number of microstates with a given energy E) is defined as

$$\Omega(E) \sim \int \mathrm{d}^{3N}q\mathrm{d}^{3N}p\delta(E - H(p, q)), \tag{6}$$

where (p, q) are the phase space coordinates, H is the Hamiltonian and we forget for the moment about multiplicative N-dependent factors and dimensional constants (for a more precise definition see Section 4). The entropy is defined via the classical Boltzmann formula

$$S(E) = \ln \Omega(E), \tag{7}$$

where we adopt units for which the Boltzmann constant k_B is equal to 1.

The non additivity has strong consequences on the construction of the canonical ensemble from the microcanonical. The reasoning usually goes as follows. One considers an isolated system with energy E, that we divide into a "small" part with energy E_1 (the subsystem of interest) and a "large" part with energy E_2 which plays the role of the bath. The additivity of the energy implies that the probability distribution that the "small" system has an energy E_1, let us call it $p(E_1)$, is given by

$$p(E_1) = \int \Omega_2(E_2)\delta(E_1 + E_2 - E)\mathrm{d}E_2 \tag{8}$$

$$= \Omega_2(E - E_1). \tag{9}$$

Using the entropy to express Ω_2 and expanding the term $S_2(E - E_1)$, one gets

$$p(E_1) = \exp\left[S_2(E - E_1)\right] \tag{10}$$

$$\approx \exp\left[S_2(E) - E_1 \left.\frac{\partial S_2}{\partial E}\right|_E + \cdots\right] \tag{11}$$

$$\propto \Omega_2(E)\mathrm{e}^{-\beta E_1}, \tag{12}$$

where

$$\beta = \left.\frac{\partial S_2}{\partial E}\right|_E. \tag{13}$$

One ends up with the usual canonical distribution for the system of interest after performing the thermodynamic limit [29]. Let us however remark that this derivation is valid also before taking the thermodynamic limit. This has led many authors to develop a thermodynamic formalism for finite systems [9].

It is clear that additivity is crucial to justify the factorization hypothesis implied in (8), and hence the existence of the canonical distribution. This hypothesis is clearly violated when the system size is finite, because of a contribution to the entropy coming from the surface that separates the two subsystems. However, after performing the thermodynamic limit, this contribution becomes negligible when the interactions are short-range. This is not true for long-range interactions which are also non-additive in the thermodynamic limit. This has led the community to split into two different attitudes. On one side are those who think that the canonical ensemble cannot be appropriately defined for long-range interactions and claim that all the analyses should be performed using the microcanonical ensemble [19,31]. Although clearly appropriate for an isolated system, this approach cannot describe open systems. On the other side are those who stress that the canonical ensemble could be still formally defined and used [3,32–34]. Among the two sides are those who attempt operative definitions of a heat bath. For instance, one has imagined that the heat bath is a short-range system that interacts with the system of interest with short-range interactions, so that the non-additivity is due only to the system [35,36]. The failure of the usual derivation of the canonical ensemble suggests that non-additive systems might have a very peculiar behavior if they are in contact with a thermal reservoir (see Refs. [37–41] for recent literature on this topic). A third position was recently initiated by Bouchet and Barré who proposed that when considering systems with long-range interaction, the canonical ensemble does not describe fluctuations of a small part of the whole system. However, they argued [42,43] that it may describe fluctuations of the whole system when coupled to a thermostat via a negligibly small coupling. This interesting line of thought needs to be pursued theoretically and confirmed by numerical simulations.

In the following section, we will discuss some physical examples before proceeding to the core of the review.

3. Physical examples of long-range interacting systems

There are many systems in nature where particles interact with a pair potential that decays at large distances as $V(r) \sim r^{-\alpha}$ with $\alpha \leq d$. Although these systems are deeply studied in their own sake (e.g. gravitational many body systems, Coulomb systems, systems of vortices, etc.), they rarely appear in books of statistical mechanics. This is of course due to the

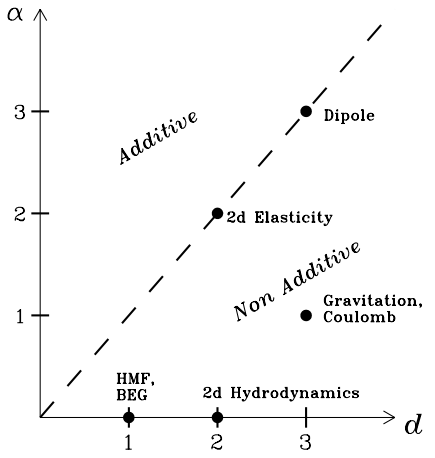


Fig. 4. Location of some physical systems in the plane where the abscissa is space dimension d and the ordinate is the exponent α characterizing the spatial decay at large distances of the pair potential.

difficulty to deal with systems that are non-extensive and non-additive. We give here a brief sketch of the physics involved in some of these systems.

In Fig. 4, we draw in the (α, d) plane some of the physical systems with long-range interactions that we will discuss afterwards.

3.1. Gravitational systems

Gravitational systems, which correspond to $\alpha = 1$ in dimension $d = 3$, clearly belong to the category of long-range interacting systems. The gravitational problem is particularly difficult because, in addition to the non-additivity due to the long-range character of the interaction, one also needs a careful regularization of the potential at short distances to avoid collapse. To be more specific, let us consider the canonical partition function of a system of N self-gravitating particles of the same mass m moving inside a volume V

$$Z_N = \frac{1}{(2\pi\lambda^2)^{3N/2}N!}\int_V \prod_{i=1}^{N} \mathrm{d}\overrightarrow{r_i}\, \exp\left[-\beta U(\overrightarrow{r}_1, \ldots, \overrightarrow{r}_N)\right], \tag{14}$$

where

$$U(\overrightarrow{r}_1, \ldots, \overrightarrow{r}_N) = -\mathcal{G}m^2 \sum_{i<j}^{N} V(|\overrightarrow{r_i} - \overrightarrow{r_j}|), \tag{15}$$

with

$$V(r) = \frac{1}{r}, \tag{16}$$

where β is the inverse temperature, $\mathcal{G}$ the gravitational constant, and $\lambda = \hbar(\beta/m)^{1/2}$ the De Broglie wavelength. In the following, we will not introduce dimensional constants, i.e. $\hbar$, in the definition of partition functions for a matter of convenience (see e.g. Eq. (55)). From the shape of the potential represented in Fig. 5 by the solid line, one clearly sees that Z_N will diverge if at least two particles collapse towards the same point. This difficulty arises because the potential is not bounded from below as for the Lennard–Jones or Morse potential. In quantum mechanics, the collapse of self-gravitating fermions [4,44], a system which is physically relevant for dwarfs and neutron stars, is forbidden by the Pauli exclusion principle, that introduces a natural effective small scale cut-off. However, to avoid the use of quantum concepts and stick to a classical model, the usual trick is to introduce an *ad-hoc* cut-off [22]. One way of regularizing the potential is shown in Fig. 5 by the dotted line. One can imagine that the potential has a hard core, which represents the particles' size. As a consequence, the inequality $U(r) \geq -Gm^2C \equiv -C'$ allows one to easily determine a finite upper bound of the configurational partition function

$$Z_N \leq \frac{V^N}{(2\pi\lambda^2)^{3N/2}N!}\exp[\beta C' N(N-1)/2]. \tag{17}$$

As far as the microcanonical ensemble is concerned, a standard argument tells us that the entropy of a self-gravitating system in a finite volume might have convergence problems if the potential is not regularized at short distances. Indeed,

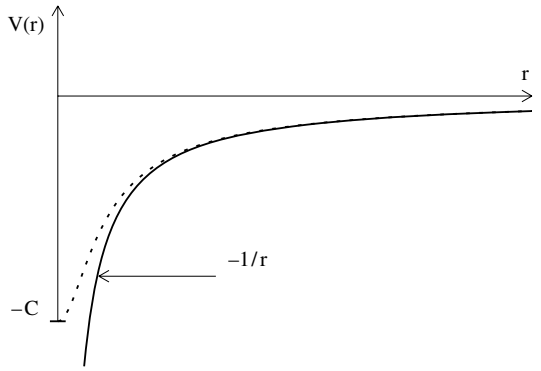


Fig. 5. The gravitational potential as a function of the distance r is represented by the solid curve. The dotted curve shows the short-distance regularized potential which avoids gravitational collapse, $-C$ is the lower bound of the potential.

let us consider N gravitating particles grouped together in a finite volume. A strong decrease of the potential energy of a pair of particles $-\mathcal{G}m^2/(|\overrightarrow{r_i} - \overrightarrow{r_j}|)$ is obtained when $\overrightarrow{r_i}$ tends to $\overrightarrow{r_j}$. Since the total energy is a conserved quantity, the kinetic energy correspondingly increases. This process leads to an increase of the accessible phase volume in the direction of momentum, and hence to an entropy increase. Since the process extends to the limit of zero distance among the particles, it might induce a divergence of the density of states, and then of Boltzmann entropy. However, one can show that this does not happen for $N = 2$ and a more subtle derivation [3,45] reveals that the entropy integral diverges only when $N \geq 3$. It is remarkable that this also corresponds to the transition from an integrable ($N = 2$) to a non-integrable ($N = 3$) gravitational system. As a consequence of this discussion, "equilibrium" states can exist only in association with local entropy maxima [17,23].

Besides these unusual properties of the canonical and microcanonical partition functions, self-gravitating systems were historically the first physical system for which ensemble inequivalence was discovered through the phenomenon of negative specific heat. The possibility of finding a negative specific heat in gravitational systems was already emphasized by Emden [15] and Eddington [16]. The divergence of the phase space volume for gravitational systems was proved by Antonov [17], and the fact that this implies negative specific heats was later stressed by Lynden-Bell [18]. An early remark on the possibility of having a negative specific heat can be found in the seminal review paper on statistical mechanics by Maxwell [46]. This was for a long time considered a paradox, until Thirring [11,12] finally clarified the controversial point by showing that the paradox disappears if one realizes that the microcanonical specific heat can be negative only in the microcanonical ensemble. Therefore, one can attribute to Thirring the discovery of ensemble inequivalence.

Let us rephrase Thirring's argument. In the canonical ensemble, the mean value of the energy is computed from the partition function (14) as

$$\langle E\rangle = -\frac{\partial \ln Z}{\partial \beta}. \tag{18}$$

It is then straightforward to compute the heat capacity at constant volume

$$C_V = \frac{\partial \langle E\rangle}{\partial T} = \beta^2 \langle (E - \langle E\rangle)^2\rangle > 0. \tag{19}$$

Let us recall that $k_B = 1$. This clearly shows that the canonical specific heat is always positive. Notice also that this condition is true for systems of any size, regardless of whether a proper thermodynamic limit exists or not.

For self-gravitating systems at constant energy (i.e., in the microcanonical ensemble) a simple physical argument which justifies the presence of a negative specific heat has been given by Lynden-Bell [47]. It is based on the virial theorem, which, for the gravitational potential, states that

$$2\langle K\rangle + \langle U\rangle = 0, \tag{20}$$

where K and U are the kinetic and potential energy, respectively. Recalling that the total energy E is constant

$$E = \langle K\rangle + \langle U\rangle = -\langle K\rangle, \tag{21}$$

where in the second identity we have used the virial theorem (20), and since the kinetic energy K defines the temperature, one gets

$$C_V = \frac{\partial E}{\partial T} \propto \frac{\partial E}{\partial K} < 0. \tag{22}$$

Losing its energy, the system becomes hotter.

There is a further difficulty in the case of gravitational interactions: the system is open, i.e. without boundary, strictly speaking. Therefore, the microcanonical partition function (6) will diverge. This divergence is actually not peculiar to self-gravitating systems since it would also occur for a perfect gas. Although for gases it is natural to confine them in a box, this is completely unjustified for gravitational systems. A way out from this pitfall is to consider an expanding universe and reconsider the problem in a wider context [33].

A detailed discussion on phase transitions in self-gravitating systems, in both the canonical and microcanonical ensemble, is not the aim of this review, and can be found in Refs. [3,4].

3.2. Two-dimensional hydrodynamics

Two-dimensional incompressible hydrodynamics is another important case where long-range interactions appear. Although high Reynolds number flows have a very large number of degrees of freedom, one can often identify structure in the flow. This suggests that one could use a much smaller number of effective degrees of freedom to characterize the flow. This remark is particularly valid for two-dimensional flows, where the inverse energy cascade leads to the irreversible formation of large coherent structures (e.g. vortices). A system with a large number of degrees of freedom which can be characterized by a small number of effective parameters, is reminiscent of what happens in thermodynamics [48], where a few macroscopic variables describe the behavior of systems composed of many particles. Statistical mechanics for turbulence is nowadays a very active field of research [49], initiated long ago by Lars Onsager [50]. We will discuss in the following the long-range character of two-dimensional hydrodynamics.

The velocity of a two-dimensional flow can be expressed in terms of the stream function $\psi(\vec{r})$, where $\vec{r} = (x, y)$ is the coordinate on the plane

$$v_x = +\frac{\partial \psi}{\partial y} \tag{23}$$

$$v_y = -\frac{\partial \psi}{\partial x}. \tag{24}$$

The vorticity ω is related to the velocity field

$$\omega = \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \tag{25}$$

and, hence, to the stream function by the Poisson equation

$$\omega = -\Delta\psi. \tag{26}$$

Using the Green's function $G\left(\vec{r}, \vec{r}\,'\right)$, one easily finds the solution of the Poisson equation in a given domain D

$$\psi(\vec{r}) = \int_D \mathrm{d}\vec{r}\,'\omega(\vec{r}\,')G\left(\vec{r}, \vec{r}\,'\right), \tag{27}$$

plus surface terms [51]. In an infinite domain,

$$G\left(\vec{r}, \vec{r}\,'\right) = -\frac{1}{2\pi}\ln|\vec{r} - \vec{r}\,'|. \tag{28}$$

The energy is conserved for the Euler equation and is given by

$$H = \int_D \mathrm{d}\vec{r}\,\frac{1}{2}(v_x^2 + v_y^2) \tag{29}$$

$$= \int_D \mathrm{d}\vec{r}\,\frac{1}{2}(\nabla\psi)^2 \tag{30}$$

$$= \frac{1}{2}\int_D \mathrm{d}\vec{r}\,\omega(\vec{r})\psi(\vec{r}) \tag{31}$$

$$= -\frac{1}{4\pi}\int_D\int_D \mathrm{d}\vec{r}\,\mathrm{d}\vec{r}\,'\omega(\vec{r}\,')\omega(\vec{r})\ln|\vec{r} - \vec{r}\,'|. \tag{32}$$

This emphasizes that one gets a logarithmic interaction between vorticies at distant locations, which corresponds to a decay with an effective exponent $\alpha = 0$, well within the case of long-range interactions (see Fig. 4). For a finite domain D the Green's function contains additional surface terms [51], which however gives no contribution to the energy (29) if the velocity field is tangent to the boundary of the domain (no outflow or inflow).

Another important conserved quantity is enstrophy, defined as

$$\mathcal{A} = \frac{1}{2}\int_D \mathrm{d}\vec{r}\,\left[\omega(\vec{r})\right]^2. \tag{33}$$

The long-range character of the interaction is even more evident if one approximates the vorticity field by N point vortices located at $\vec{r}_i = (x_i, y_i)$, with a given circulation Γ_i

$$\omega(\vec{r}) = \sum_{i=1}^{N} \Gamma_i \delta(\vec{r} - \vec{r}_i). \tag{34}$$

The energy of the system now reads

$$H = -\frac{1}{4\pi}\sum_{i\neq j} \Gamma_i\Gamma_j \ln|\vec{r}_i - \vec{r}_j|, \tag{35}$$

where we have dropped the self-energy term because, although singular, it would not induce any motion [52]. Considering the two coordinates of the point vortex on the plane, the equations of motion are

$$\Gamma_i \frac{\mathrm{d}x_i}{\mathrm{d}t} = +\frac{\partial H}{\partial y_i} \tag{36}$$

$$\Gamma_i \frac{\mathrm{d}y_i}{\mathrm{d}t} = -\frac{\partial H}{\partial x_i}. \tag{37}$$

The phase-space volume contained inside the energy shell $H = E$ can be written as

$$\Phi(E) = \int \prod_{i=1}^{N} \mathrm{d}\vec{r}_i\,\theta(E - H(\vec{r}_1, \ldots, \vec{r}_N)), \tag{38}$$

θ being the Heaviside step function. The total phase space volume is $\Phi(\infty) = A^N$, where A is the area of the domain D. One immediately realizes that $\Phi(E)$ is a non-negative increasing function of the energy E with limits $\Phi(-\infty) = 0$ and $\Phi(\infty) = A^N$. Therefore, its derivative, which is nothing but the microcanonical partition function (6), is given by

$$\Omega(E) = \Phi'(E) = \int \prod_{i=1}^{N} \mathrm{d}\vec{r}_i\,\delta(E - H(\vec{r}_1, \ldots, \vec{r}_N)), \tag{39}$$

and is a non-negative function going to zero at both extremes $\Omega(\pm\infty) = 0$. Thus the function must achieve at least one maximum at some finite value E_m where $\Omega'(E_m) = 0$. For energies $E > E_m$, $\Omega'(E)$ will then be negative. Using the entropy $S(E) = \ln\Omega(E)$, one thus gets that the inverse temperature $\mathrm{d}S/\mathrm{d}E$ is negative for $E > E_m$. This argument for the existence of negative temperatures was proposed by Onsager [50] two years before the experiment on nuclear spin systems by Purcell and Pound [53] reported the presence of negative "spin temperatures".

Onsager also pointed out that negative temperatures could lead to the formation of large-scale vortices by a clustering of smaller ones. Although, as anticipated, the canonical distribution has to be used with caution for long-range interacting systems, the statistical tendency of vortices of the same circulation sign to cluster in the negative temperature regime can be justified using the canonical distribution $\exp(-\beta H)$. Changing the sign of the inverse temperatures β would be equivalent to reversing the sign of the interaction between vortices, making them repel (resp. attract) if they are of opposite (resp. same) circulation sign.

After a long period in which Onsager's statistical theory was not further explored, this domain of research has made an impressive progress recently. One might in particular cite the work of Joyce and Montgomery [54], who have considered a system of vortices with total zero circulation. Maximizing the entropy at fixed energy, they obtain an equation for the stream function which gives exact stable stationary solutions of the 2D Euler equations, able to describe the macroscopic vortex formation proposed by Onsager for negative temperatures. Later, Lundgren and Pointin [55] studied the effect of far vorticity field on the motion of a single vortex, showing that it produces a positive eddy viscosity term leading to an increase of cluster size. Subsequently, Robert [6] and Miller [7] have elaborated an equilibrium statistical mechanical theory directly for the continuum 2D Euler equation. Nice prolongations along these lines are Refs. [24,56,57], together with applications to the Great Red Spot of Jupiter [58–61].

The canonical ensemble can also be defined for both the Euler equation in 2D and the Onsager point vortex model. It turns out that in both cases, canonical and microcanical ensembles may be inequivalent [13,62–64]. In particular for the Euler equation in the region of negative temperature (where vorticity tends to accumulate at the center of the domain) some hybrid states are shown to be realized in the microcanonical ensemble but not in the canonical [63]. As far as the point vortex model is concerned, microcanonical stable states, that are unstable in the canonical ensemble, are found for specific geometries [13]. In both cases, the microcanonical specific heat is negative. Reviews on 2D turbulence can be found in Refs. [34,65–67]. Another related interesting case of ensemble inequivalence has been recently reported in the context of physical oceanography [68].

3.3. Two-dimensional elasticity

Let us discuss the planar stress and displacement fields around the tip of a slit-like plane crack in an ideal Hookean continuum solid. The classical approach to a linear elasticity problem of this sort involves the search for a suitable "stress function" that satisfies the so-called biharmonic equation

$$\nabla^2(\nabla^2\psi) = 0 \tag{40}$$

where ψ is the Airy stress function, which has to satisfy appropriate boundary conditions. The deformation energy density is then defined as $U \propto \sigma\epsilon$ where σ is the fracture stress field around the tip, whereas ϵ is the deformation field. Considering a crack-width a and using the exact Muskhelishvili's solution [69], one obtains the elastic potential energy due to the crack

$$U \simeq \frac{\sigma_\infty^2(1-\nu)}{2E}\frac{a^2}{r^2}, \tag{41}$$

where E is the Young modulus, σ_∞ the stress field at infinity, ν the Poisson coefficient and r the distance to the tip. The elasticity equation in the bulk of two-dimensional materials leads therefore to a marginal case of long-range interaction, since $U \sim 1/r^2$ in $d = 2$. Looking at the important engineering applications, the dynamics of this non-conservative system should be better studied: the difficulty lies again in the long-range nature of the interaction. In addition, in such a two-dimensional material, the presence of several fractures could exhibit a very interesting new type of screening effects.

3.4. Charged systems

The partition function of system of charges is basically the same as the one for gravitational systems displayed in formula (14) with the potential given by

$$U(\vec{r}_1, \ldots, \vec{r}_N) = \frac{1}{4\pi\varepsilon_0}\sum_{i<j}^{N} e_i e_j V(|\vec{r_i} - \vec{r_j}|), \tag{42}$$

where $e_i = \pm e$ is the charge of the particle and V is given by formula (16). If the total charge is non-zero, the excess charge is expelled to the boundary of the domain and the bulk is neutral. Hence, neglecting boundary contributions (which are non-extensive), one usually considers an infinite medium with total zero charge [70,71]. Similarly to the gravitational case, the partition function would diverge if not properly regularized at short distances. This is usually done by supposing that additional forces are present at short distances, either hard core of radius λ or smoothed Coulombic singularities of the type

$$V_{smooth} \sim \frac{1-\exp(-r/\lambda)}{r}. \tag{43}$$

The regularized partition function allows one to perform the thermodynamic limit and derive physical quantities, like the pressure.

As far as the large distance behavior is concerned, several rigorous results exist [72] that prove, under appropriate hypotheses (low density and high temperature), that the effective two-body potential is Debye–Hückel screened

$$V_{eff} \propto \frac{\exp(-r/\ell_D)}{r}, \tag{44}$$

where $\ell_D = (\varepsilon_0/(2ne^2\beta))^{1/2}$ is the Debye length, with n the density. Among the hypotheses, the most important one for its possible physical consequences is the one of low density or high temperature. This suggests that all pathologies related to ensemble inequivalence (e.g. negative specific heat), that is a consequence of the long-range nature of the interaction, will be indeed absent for charged systems.

An interesting different situation is the one of plasmas consisting exclusively of single charged particles (pure electron or pure ion plasmas) [73]. Charged particles are confined by external electric and magnetic fields. The thermodynamics of such systems is not affected by short-distance effects because particles repel each other. On the contrary, the behavior at large distances is different from the one of globally neutral low density plasmas. Nevertheless, by studying the correlation at equilibrium, a Debye length emerges [73] (this behavior is also present in self-gravitating systems). Experimentally, relaxation to thermal equilibrium has been observed only in some specific conditions. The system has been shown to relax into quasi-stationary states, including minimum enstrophy states (see Eq. (33)) and vortex crystal states [74,75]. Indeed, in the particular case of a pure electron plasma in a cylindrical container with a strong magnetic field applied along the axis of the cylinder, it can be shown that the electron motion in the plane perpendicular to the magnetic field obeys equations that are the same as the Onsager point vortex model (36) and (37). Therefore, all of what has been written above about two-dimensional hydrodynamics applies to this system, including the existence of negative specific heat which has been explicitly demonstrated in a magnetically self-confined plasma torus [76].

3.5. Dipolar systems

Systems of electric and magnetic dipoles share many similarities, but also have important differences (force in non-uniform external fields, the symmetry axial/polar of the dipolar vector). However, for what concerns statistical mechanics, the two systems are equivalent. Magnetic dipolar systems are easier to realize in nature; let us then concentrate on them. The interaction energy between two magnetic dipoles is

$$E_{ij} = \frac{\mu_0}{4\pi}\left[\frac{\vec{\mu}_i \cdot \vec{\mu}_j}{|\vec{r}_{ij}|^3} - \frac{3(\vec{\mu}_i \cdot \vec{r}_{ij})(\vec{\mu}_j \cdot \vec{r}_{ij})}{|\vec{r}_{ij}|^5}\right], \tag{45}$$

where $\vec{\mu}_i$ is the magnetic moment, $\vec{r}_{ij} = (\vec{r}_j - \vec{r}_i)$ is the distance between the two dipoles and μ_0 is the magnetic permeability of the vacuum. Dipolar interaction energy is strongly anisotropic: on a lattice, magnetic moments parallel with a bond interact ferromagnetically, while when they are perpendicular to the bond they interact antiferromagnetically. When magnetic moments are placed on a triangular or a square lattice the interaction is *frustrated*. Dipolar forces are long-range only in $d = 3$, because the energy per spin scales as $\int \mathrm{d}^d r/r^3$. They are therefore *marginally* long-range in $d = 3$, while they are short-range in $d = 1, 2$. It should be remarked that elasticity is instead marginally long-range in $d = 2$.

It can be shown in general that, for samples of ellipsoidal shape and when the magnetization $\vec{M} = \sum_i \vec{\mu}_i/V$ lies along the longest principal axis of the ellipsoid, the energy per volume of a system of dipoles on a lattice can be written as

$$\frac{H}{V} = \frac{1}{2V}\sum_{i,j} E_{ij} = E_0 + \frac{1}{2}\mu_0 |M|^2 D, \tag{46}$$

where E_0 is an energy that depends on the crystal structure, and D is the so called demagnetizing factor, which is equal to $1/3$ for spherically shaped samples, tends to $D = 0$ for needle shape samples and to $D = 1$ for disk shaped ones. The well-known *shape dependence* of dipolar energy is hence accounted for by the highly frustrating antiferromagnetic demagnetizing term in formula (46) [77]. We will see examples of energies of the form (46) in Section 6.1, while in Section 6.2.4, we will work out in detail a specific example corresponding to a realistic system.

Using methods similar to those introduced to prove the existence of a thermodynamic limit for short-range forces [29], it can be shown [78,79] that a system of dipolar spins posses a well defined bulk free energy, independent of sample shape, only in the case of zero applied field. The key to the existence of this thermodynamic limit is the reduction in demagnetization energy when uniformly magnetized regions break into ferromagnetically ordered domains [80]. Technically, the proof is performed by showing that under the hypothesis of zero field the free energy is additive, using invariance under time reversal of the dipole energy, which is a consequence of its bilinearity in $\vec{\mu}_i$.

In the future, magnetic dipolar systems might constitute a field where the results on the statistical mechanics of systems with long-range will find a fruitful application. Besides the example discussed in Section 6.2.4, holmium titanate materials [81], where dipolar interactions dominate over Heisenberg exchange energy, deserve some attention. The possibility to perform experiments with single domain needle shaped dipolar materials has been also stressed [82].

3.6. Small systems

As we have seen, the presence of long-range interactions causes the lack of additivity in macroscopic systems. However, even if the interactions are short-range, systems of a linear size comparable to the range of the interaction are non-additive. Then, we should expect that some of the peculiar features found for macroscopic systems with long-range interactions are present also in microscopic or mesoscopic systems. Examples of this sort are: atomic clusters, quantum fluids, large nuclei, dense hadronic matter.

As it will be shown in the following section, the study of phase transitions is very important for the characterization of the properties of macroscopic systems with long-range interactions, especially in relation to the issue of ensemble inequivalence. However, phase transitions do occur also in atomic clusters (liquid–gas and solid–liquid transitions), quantum fluids (Bose–Einstein condensation or super-fluid transition), large nuclei (liquid–gas transition), dense hadronic matter (formation of quark-gluon plasma). Extensive studies have been devoted to phase transitions in the thermodynamic limit. This limit introduces simplifications in the analytical treatment, also in the case of long-range interactions. On the contrary, a consistent theory of phase transitions for small systems has not yet been developed. Signatures of phase transitions in finite systems are, however, often found both in numerical and laboratory experiments.

Gross has focused his attention on the theoretical treatment of "Small" systems [19,31,83]. His approach privileged the use of the microcanonical ensemble. Therefore, he immediately realized the possibility that specific heat could be negative and pointed out the feasibility of experiments with heavy nuclei.

According to Chomaz and collaborators, the key point is to begin with the general definition of entropy in the framework of information theory (see Refs. [9,84–88]). In order to treat on the same ground classical and quantum systems, Chomaz introduces the density matrix

$$\hat{D} = \sum_n |\Psi_n\rangle\, p_n\, \langle\Psi_n|\,, \tag{47}$$

where $|\Psi_n\rangle$ are the states of the system and p_n their probabilities. The entropy is thus defined by

$$S[\hat{D}] = -\mathrm{Tr}\,\hat{D}\ln\hat{D}. \tag{48}$$

Different Gibbs ensembles are obtained by maximizing $S[\hat{D}]$ with respect to the probabilities p_n under some constraints. Each set of constraints defines an ensemble. The different number and functional forms of the constraints are related to the different physical situations. Within this framework, the finiteness of small systems is described by the introduction of specific constraints. It has been shown that phase transitions in finite systems can be equivalently signaled by three different effects. The first one is the bimodality of the density of states as a function of energy [89], with the distance between maxima corresponding to different phases scaling as the number of particles. The second one is a negative slope of the microcanonical caloric curve, i.e. a negative specific heat [90]. The third one is the presence of anomalously large fluctuations in the energy partition between potential energy and kinetic energy. When the interactions are short-range, all these features give rise in the thermodynamic limit to the usual phase transitions, with the disappearance of the negative slope of the caloric curve. It is also possible to find a further signature of a phase transition, that makes a connection with Yang–Lee theory of phase transitions. In this latter theory, all zeroes of the partition function lie in the complex plane of the temperature for the canonical ensemble or of the fugacity for the grand-canonical ensemble, with an imaginary part different from zero as long as the system is finite. Phase transitions in the infinite system are associated to the approach of some of these zeros to the real axis, as system size increases. It has been stressed that the way in which zeroes approach the real axis may serve as a classification of phase transitions in finite systems [87].

Some of the previously mentioned signatures of phase transitions in finite systems have been also experimentally reported. We mention here experiments on atomic clusters [91,92], and experiments on nuclear fragmentation [93]. In all experiments, the microcanonical caloric curve is compatible with the presence of an energy range where the specific heat is negative.

The first set of experiments is realized using atomic sodium clusters Na^+_{147} and hydrogen cluster ions $H_3^+(H_2)_{m\leq 14}$. In the first case, the negative specific heat has been found in correspondence to a solid–liquid phase transition, while in the second case in the vicinity of a liquid–gas transition. Sodium clusters [91] are produced in a gas aggregation source and then thermalized with helium gas at a controlled temperature and selected to a single cluster size by a first mass spectrometer. Energy of the clusters is then increased by laser irradiation leading finally to evaporation. A second mass spectrometer allows the reconstruction of the size distribution, and correspondingly of their energies. Performing this experiment at different temperatures of the helium gas, a caloric curve is constructed. The procedure assumes that, after leaving the source, a microcanonical temperature can be assigned to the clusters. Conceptually, this is probably the most delicate point of the experiment. A region of negative specific heat, corresponding to the solid–liquid transition, has been reported.

In the second set of experiments, performed with hydrogen cluster ions [92], the energy and the temperature are determined from the size distribution of the fragments after collision of the cluster with a Helium projectile. This is done using a method introduced in Ref. [94]. The reported caloric curve [92] shows a plateau. Work along this line is in progress and seems to show a negative specific heat region [95], corresponding to a liquid–gas transition.

In the third set of experiments on nuclear fragmentation [93], the presence of negative specific heat is inferred from the event by event study of energy fluctuations in excited Au nuclei resulting from Au + Au collisions. The data seem to indicate a negative specific heat at an excitation energy around 4.5 MeV/u. However, the signature corresponds to indirect measurements, and the authors cautiously use the word "indication" of negative specific heat.

4. Equilibrium statistical mechanics: Ensemble inequivalence

Our purpose in this section is to propose a fully consistent statistical mechanics treatment of systems with long-range interactions. An overview of the different methods (saddle-point techniques, large deviations, etc.) will be presented. Simple mean-field models will be used both for illustrative purposes and as concrete examples. Most of the features of long-range systems that we will discuss are valid also for non-mean-field Hamiltonians as will be discussed in Section 6.

4.1. Equivalence and inequivalence of statistical ensembles: Physical and mathematical aspects

4.1.1. Ensemble equivalence in short-range systems

In short-range systems, statistical ensembles are equivalent. A thorough proof of this result can be found in the book by Ruelle [29]. Let us first discuss the physical meaning of ensemble equivalence in order to clarify why for long-range interacting systems equivalence does not always hold.

The three main statistical ensembles are associated to the following different physical situations:

- a completely isolated system at a given energy E: *microcanonical* ensemble;
- a system that can exchange energy with a large thermal reservoir characterized by the temperature T: *canonical* ensemble;
- a system that can exchange energy and particles with a reservoir characterized by the temperature T and the chemical potential μ: *grand canonical* ensemble.

Equivalence of the ensembles relies upon two important physical properties:

(i) in the thermodynamic limit, excluding critical points, the relative fluctuations of the thermodynamic parameters that are not held fixed (e.g. energy in the canonical ensemble) vanish;
(ii) a macroscopic physical state that is realizable in one ensemble can be realized also in another (equivalence at the level of macrostates).

Let us concentrate our attention on the second item and let us refer to the equivalence between the microcanonical and the canonical ensemble: an isolated system with a given energy has an average temperature, i.e., an average kinetic energy. If instead we put the system in contact with a thermal bath at that temperature, we have an average energy equal to the energy of the isolated system. Therefore there is a one-to-one correspondence between energy values and temperature values. Actually, in the presence of phase transitions, this statement has to be made more precise, as we will comment in Section 4.1.2.

The practical consequence of ensemble equivalence is that, for computational purposes, one has the freedom to choose the ensemble where calculations are easier, and typically this is not the microcanonical ensemble (it is easier to integrate Boltzmann factors than δ-functions). Thus, in spite of its fundamental importance in the construction of statistical mechanics, the microcanonical ensemble is practically never used to perform analytical calculations. On the contrary it is very much used in numerical simulations, since it constitutes the fundamental ingredient of molecular dynamics [96].

Ensemble equivalence is mathematically based on certain properties of the partition functions. To illustrate this point, we again consider the microcanonical and canonical ensembles, referring the reader to Ruelle [29] for a complete and rigorous discussion. A more precise definition of the microcanonical partition function (see formula (6)) of a system in $d = 3$ with N particles confined in a volume V is given by

$$\Omega(E, V, N) = \frac{1}{N!} \int_{\Gamma} \mathrm{d}q^{3N} \mathrm{d}p^{3N} \delta\left(E - H(p, q)\right), \tag{49}$$

where the domain of integration is the accessible phase space Γ.

For lattice systems the definition of Ω is slightly different. There is no explicit volume dependence (because volume is fixed once the lattice constant and N are given) and no $N!$ term due to the distinguishability of the lattice sites (for more details on the Gibbs paradox see [10]). The microcanonical partition function is in this case

$$\Omega_{lattice}(E, N) = \int \prod_{i}^{N} \mathrm{d}\mathbf{q}_i \prod_{i}^{N} \mathrm{d}\mathbf{p}_i \delta\left(E - K(\{\mathbf{p}_i\}) - U(\{\mathbf{q}_i\})\right), \tag{50}$$

where $\mathbf{q}_i$ and $\mathbf{p}_i$ are the conjugate variables attached to site i (see formula (5)) and K the kinetic energy.

The entropy is defined by

$$S(E, V, N) = \ln \Omega(E, V, N). \tag{51}$$

The thermodynamic limit corresponds to $N \to \infty$, $E \to \infty$ and $V \to \infty$ such that $N/V \to n$ and $E/N \to \varepsilon$, where the density $n \geq 0$ and the energy per particle ε are finite. The limit

$$s(\varepsilon, n) = \lim_{N\to\infty} \frac{1}{N} S(E, V, N) \tag{52}$$

exists and gives the entropy per particle. The function $s(\varepsilon, n)$ is continuous, increasing in ε at fixed n, so that the temperature

$$T = (\partial s/\partial \varepsilon)^{-1} \tag{53}$$

is positive. Measuring the temperature by this formula seems hardly feasible. However, for Hamiltonians with kinetic energy, it can be shown that T coincides with the average kinetic energy, which is accessible experimentally [97]. For short-range systems, $s(\varepsilon, n)$ is a concave function of ε at fixed n, i.e.

$$s\left(c\varepsilon_1 + (1 - c)\varepsilon_2, n\right) \geq c\, s(\varepsilon_1, n) + (1 - c)\, s(\varepsilon_2, n) \tag{54}$$

for any choice of ε_1 and ε_2, with $0 \leq c \leq 1$ (in lattice systems without kinetic energy the energy can be bounded from above, and in turn this implies that T can be negative; however, concavity is still guaranteed if the interactions are short-range). This property is important in connection with the partition function of the canonical ensemble, given by

$$Z(\beta, V, N) = \frac{1}{N!} \int_{\Gamma} \mathrm{d}q^{3N} \mathrm{d}p^{3N} \exp\left[-\beta H(p, q)\right], \tag{55}$$

with $\beta \geq 0$ the inverse temperature. A similar definition of a canonical lattice partition function can be given, as done for the microcanonical lattice partition function (50). In the thermodynamic limit, the free energy per particle is

$$f(\beta, n) = -\frac{1}{\beta} \lim_{N\to\infty} \frac{1}{N} \ln Z(\beta, V, N). \tag{56}$$

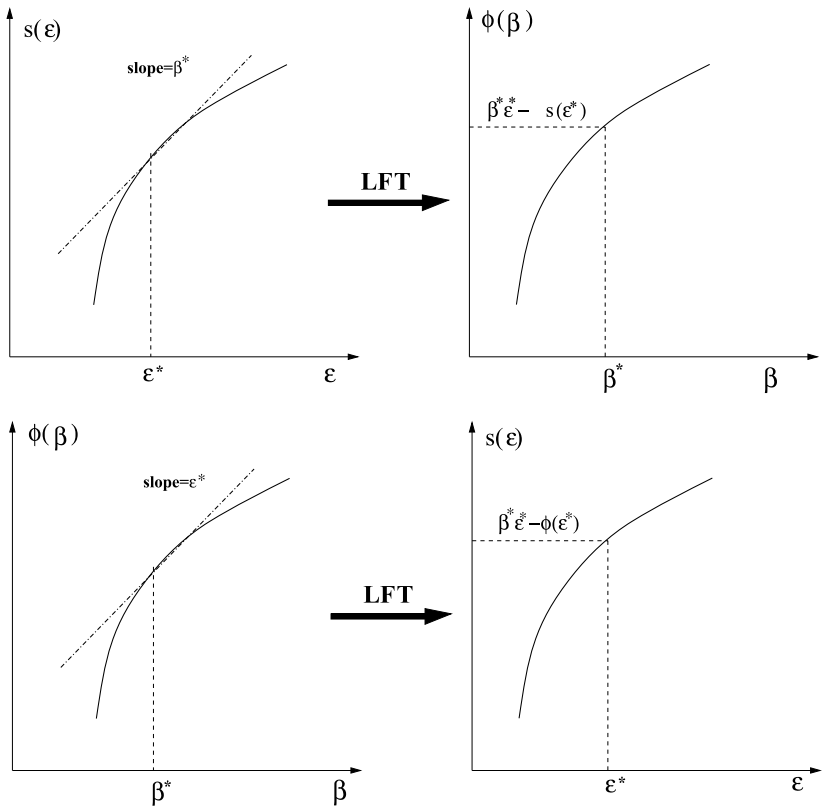


Fig. 6. Relation between the entropy per particle $s(\varepsilon, n)$ and $\phi(\beta, n) = \beta f(\beta, n)$ (where $f(\beta, n)$ is the free energy per particle) by the Legendre–Fenchel Transform (LFT), n is fixed.

Moreover, at fixed n, the function $\phi(\beta, n) \equiv \beta f(\beta, n)$ is concave in β. In the following $\phi(\beta, n)$ will be called the *rescaled free energy*.

The equivalence between the microcanonical and the canonical ensemble is a consequence of the concavity of ϕ and s and of the relation between these two functions given by the Legendre–Fenchel Transform (LFT). Indeed, one can prove that $\phi(\beta, n)$ is the LFT of $s(\varepsilon, n)$

$$\phi(\beta, n) = \beta f(\beta, n) = \inf_{\varepsilon} [\beta\varepsilon - s(\varepsilon, n)]\,. \tag{57}$$

A brief sketch of the proof goes as follows

$$\exp(-\beta N f(\beta, n)) = Z(\beta, V, N) \tag{58}$$

$$= \frac{1}{N!} \int \mathrm{d}E \int \mathrm{d}q^{3N} \mathrm{d}p^{3N}\, \delta(H(p, q) - E) \exp(-\beta E), \tag{59}$$

$$= \int \mathrm{d}E\, \Omega(E, V, N) \exp(-\beta E), \tag{60}$$

$$= \int \mathrm{d}E \exp\left(-N\left[\beta\varepsilon - s(\varepsilon, n)\right]\right), \tag{61}$$

where the last equality is valid for large N. The saddle point of the last integral gives formula (57). Let us remark that it would not have been necessary to hypothesize the concavity of ϕ in β from the beginning, because it follows from the fact that the LFT of a generic function is a concave function. Also the inverse LFT holds, since $s(\varepsilon, n)$ is concave in ε

$$s(\varepsilon, n) = \inf_{\beta} [\beta\varepsilon - \phi(\beta, n)]\,. \tag{62}$$

This indeed proves ensemble equivalence, because for each value of β there is a value of ε that satisfies Eq. (57), and, conversely, for each value of ε there is a value of β satisfying Eq. (62). Fig. 6 provides a visual explanation of the relation between s and ϕ and of the correspondence between ε and β.

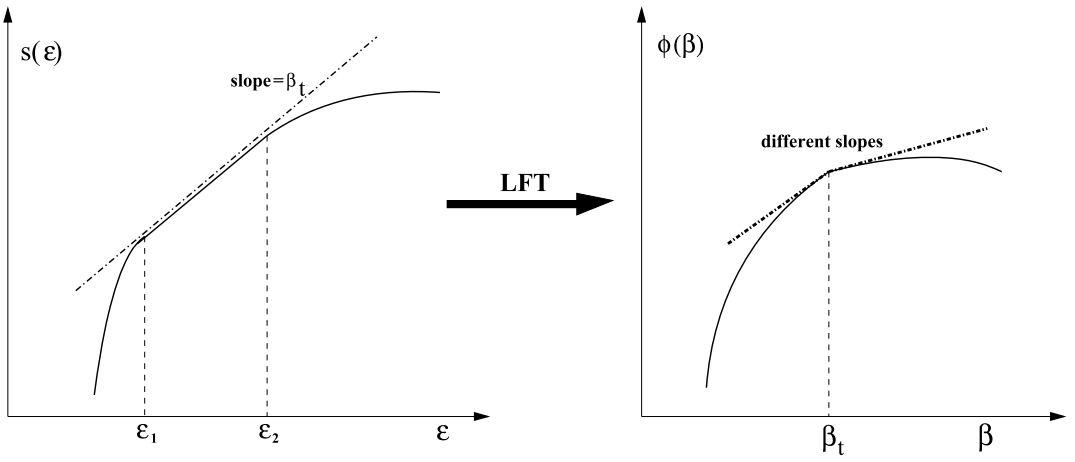


Fig. 7. Entropy $s(\varepsilon)$ and rescaled free energy $\phi(\beta)$ in the case of a first order phase transition. The inverse transition temperature is β_t and $[\varepsilon_1, \varepsilon_2]$ is the energy range of phase coexistence.

Relations similar to those described in this subsection explain the equivalence of other ensembles. For instance Van Hove [98] proved that the equation of state for a short-range classical system is the same in the canonical and grand-canonical ensemble. This implies that isothermal compressibility is positive in both ensembles.

4.1.2. Phase separation and Maxwell construction in short-range systems

It is important to discuss ensemble equivalence in the presence of phase transitions. We will see that some interesting features arise. Phase transitions are associated to singularities of thermodynamic functions [10]. Therefore, in the microcanonical and canonical ensemble, they will be signaled by discontinuities in a derivative of some order of the entropy s or the rescaled free energy ϕ.

Let us concentrate on the dependence of s and ϕ on ε and β, respectively. Consider for example a model in which, for some choice of parameters, $s(\varepsilon)$ has a zero curvature in some energy range $[\varepsilon_1, \varepsilon_2]$. At both extremes of this interval the second derivative of s has a discontinuity (see Fig. 7). Within the range $[\varepsilon_1, \varepsilon_2]$, the function is not strictly concave (i.e., Eq. (54) is satisfied with an equality). All energy values in this range correspond to the same value of the inverse temperature $\beta = \beta_t$, the slope of the straight segment of $s(\varepsilon)$ in Fig. 7. For each energy in this range, the system separates in two phases of different energies ε_1 and ε_2. Therefore, the energy will be given by

$$\varepsilon = c\varepsilon_1 + (1 - c)\varepsilon_2 \tag{63}$$

where c is the fraction of phase 1 and, of course, $1 - c$ the fraction of phase 2. We emphasize that this relation is a direct consequence of additivity. It is easy to show that the Legendre–Fenchel transform of $s(\varepsilon)$, the rescaled free energy (shown in Fig. 7) has a discontinuity in the first derivative with respect to β at β_t $((\mathrm{d}\phi/\mathrm{d}\beta)_{\mp} = \varepsilon_{1,2})$. Following Ehrenfest's classification, this is a first order phase transition.

Let us remark that in this example there is no one-to-one correspondence between ε and β: several microcanonical macroscopic states are represented by a single canonical state. This shows that, in the presence of first-order phase transitions the relation between the ensembles must be considered with care. We would not say that the ensembles are inequivalent in this case, which is a marginal one; therefore we do not adopt the term used in mathematical physics in this case: "partial equivalence" [99].

Tuning the parameters of the model in such a way that the straight segment in Fig. 7 reduces to one point, one recovers a strictly convex entropy function and a one-to-one correspondence between ε and β. Consequently, the discontinuity in the first derivative of the rescaled free energy is removed. This is the case of a second order phase transition.

It could also happen that the entropy, instead of showing the straight segment of Fig. 7, has a convex region (see the full line in Fig. 8a). For short-range interactions this is what is observed for finite systems near a phase transition. It has been shown numerically [9,19,87] and for simple models [100,101] that by increasing system size the entropy approaches the "concave envelope" which is constructed by replacing the full line in the energy range $[\varepsilon_1, \varepsilon_2]$ by the straight thick dashed line in Fig. 8a. In statistical mechanics, this procedure goes under the name of Maxwell's construction and is mostly known in connection with the Van der Waals theory of liquid–gas transition [10]. The relation between the construction of the "concave envelope" and the Maxwell construction can be easily established by looking at the following relation

$$s(\varepsilon_2) = s(\varepsilon_1) + \int_{\varepsilon_1}^{\varepsilon_2} \mathrm{d}\varepsilon \beta(\varepsilon), \tag{64}$$

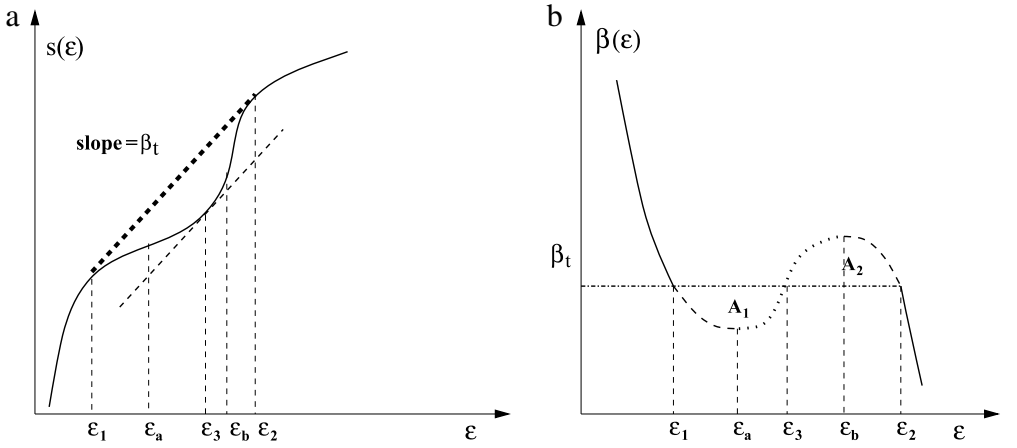


Fig. 8. (a) Schematic shape of the entropy s as a function of the energy ε (solid line) showing a "globally" convex region in the range $[\varepsilon_1, \varepsilon_2]$, the thick dashed line realizes the "concave envelope". (b) Inverse temperature β as a function of energy ε. According to Maxwell's construction $A_1 = A_2$. The curve $\beta(\varepsilon)$ represents states that are stable (solid line), unstable (dotted line) and metastable (dashed lines).

which derives from the definition of the inverse temperature $\beta = \mathrm{d}s/\mathrm{d}\varepsilon$, which is plotted in Fig. 8b as a function of energy. Another way of obtaining $s(\varepsilon_2)$ is by integrating along the thick dashed line in Fig. 8a.

$$s(\varepsilon_2) = s(\varepsilon_1) + (\varepsilon_2 - \varepsilon_1)\beta_t, \tag{65}$$

where $\beta_t = \beta(\varepsilon_1) = \beta(\varepsilon_2)$. This implies that

$$\int_{\varepsilon_1}^{\varepsilon_2} \mathrm{d}\varepsilon \beta(\varepsilon) = (\varepsilon_2 - \varepsilon_1)\beta_t. \tag{66}$$

Splitting the integral in two intervals $[\varepsilon_1, \varepsilon_3)$ and $[\varepsilon_3, \varepsilon_2]$ one gets

$$(\varepsilon_1 - \varepsilon_3)\beta_t + \int_{\varepsilon_1}^{\varepsilon_3} \mathrm{d}\beta \beta(\varepsilon) = (\varepsilon_2 - \varepsilon_3)\beta_t + \int_{\varepsilon_2}^{\varepsilon_3} \mathrm{d}\beta \beta(\varepsilon). \tag{67}$$

The value ε_3 is the one obtained from the entropy by looking where, in the convex region, the entropy has slope β_t. Condition (67) is equivalent to the equality of the areas A_1 and A_2 in Fig. 8. Introducing the *generalized free energy*, which is a function of both energy and inverse temperature,

$$\widehat{f}(\beta, \varepsilon) = \varepsilon - \frac{1}{\beta} s(\varepsilon), \tag{68}$$

one obtains from Eq. (67) that

$$\widehat{f}(\beta_t, \varepsilon_1) = \widehat{f}(\beta_t, \varepsilon_2). \tag{69}$$

This shows that the requirement that the entropy be concave is equivalent to Maxwell's equal areas construction and, in turn, equivalent to demand that the generalized free energies, computed at the transition inverse temperature β_t and at the two energies ε_1 and ε_2 which delimit the coexistence region, are equal (and looking at Fig. 8a, also equal to $f(\beta_t)$).

The Maxwell construction is related to the application of a maximum entropy principle for additive systems. Indeed, for all energies in the range $(\varepsilon_1, \varepsilon_2)$, the entropy corresponding to the full line in Fig. 8a is smaller than the entropy corresponding to the dashed line at the same energy. This latter entropy is related to a system which has performed *phase separation* and is therefore obtained as a mixture composed of a certain fraction of a state with energy ε_1 and the remaining fraction with energy ε_2, as in formula (63). This latter system having a larger entropy, the natural tendency will be to phase separate. Hence the "concave envelope" recovers maximum entropy states.

It should be remarked that the truly "locally" convex part of the entropy is the one in the range $[\varepsilon_a, \varepsilon_b]$, while the range $[\varepsilon_1, \varepsilon_2]$ is "globally" convex. One should therefore expect a difference in the properties of the physical states in the various ranges. Indeed, states in the range $[\varepsilon_a, \varepsilon_b]$ are *unstable* (dotted line in Fig. 8b), while states in the ranges $[\varepsilon_1, \varepsilon_a]$ and $[\varepsilon_b, \varepsilon_2]$ are *metastable* (dashed lines in Fig. 8b): at a solid–liquid phase transition they would correspond to superheated solids and supercooled liquids, respectively. While the unstable states cannot be observed, the metastable states are observable but are not true equilibrium states, because higher entropy phase separated states are accessible.

Considering small systems again, the nucleation of a bubble might lead to creation of a convex entropy region. Indeed, once a bubble of phase 1 is nucleated in phase 2, the energetic cost of the interface is proportional to the surface, while the energetic gain is proportional to the volume. If the system is small enough, these two energies might be comparable, implying that the additivity property is not satisfied [102,103].

4.1.3. Ensemble inequivalence in long-range systems: Negative specific heat

As already anticipated in Section 1, an important physical property of systems with long-range interactions is that ensembles can be inequivalent. This means that experiments realized in isolated systems, described by the microcanonical ensemble, may give different results from similar experiments performed with well thermalized systems, for which the canonical ensemble is the appropriate one. For instance, while the specific heat will turn out to be always positive for a system in contact with a heat bath, it might be negative for an isolated long-range system.

When the interactions are long-range, an entropy function with a convex "intruder" [19] like the one shown by the solid line in Fig. 8a can represent truly stable equilibrium states. In Section 4.2, we will give a concrete example to illustrate this important property. Here, we will develop some general considerations which are not specific to a given model.

The construction which has led to the "concave envelope" for short-range systems cannot be realized for long-range systems. On one hand the same notion of phase is ill defined for long-range systems (which are inherently inhomogeneous). On the other hand, even if a definition of phase were possible, the lack of additivity of long-range systems would not allow one to obtain a mixed state and, in particular, to derive relations like (63).

The starting point for the construction of a consistent thermodynamics of long-range systems is the calculation of microcanonical entropy associated to a given *macrostate*. A *microstate* is defined by the phase space variables of the system, and thus it refers to a precise microscopic state, while a *macrostate* is described in terms of a few macroscopic or coarse-grained variables, and then it generally defines a large set of microscopic states, all of them giving rise to the same values of the macroscopic variables. The derivation of free energy from microcanonical entropy using the Legendre–Fenchel transform (57) is still valid for long-range systems, ensuring that the function ϕ is concave also for these systems. However, when the entropy has a convex region, the inversion of the Legendre–Fenchel transform, Eq. (62) does not give the correct microcanonical entropy, but rather its "concave envelope" [99]. Physically, this implies a lack of equivalence of ensembles at the level of macrostates, i.e. all microcanonical macrostates with energies between ε_1 and ε_2 do not have a corresponding macrostate in the canonical ensemble [20].

The existence of a convex "intruder" in the entropy–energy curve, as in Fig. 8a, is associated to the presence of negative specific heat. Indeed,

$$\frac{\partial^2 S}{\partial E^2} = -\frac{1}{C_V T^2} \tag{70}$$

where the heat capacity at fixed volume is $C_V = \partial E/\partial T$. Hence, in the energy range $[\varepsilon_a, \varepsilon_b]$, the convexity of the entropy, $\partial^2 S/\partial E^2 > 0$, implies that the heat capacity is negative $C_V < 0$. This, in turn, implies that the conveniently normalized specific heat $c_V = C_V/N$ is also negative.

In the canonical ensemble, the specific heat is always positive, even if the interactions are long-range. This is a straightforward consequence of the concavity of the function ϕ, which is given by Eq. (57) also for long-range systems. Indeed,

$$\frac{\partial^2 \phi}{\partial \beta^2} = -\frac{c_V}{T^2} < 0, \tag{71}$$

implying that $c_V > 0$. There is a subtlety related to the calculation of c_V at β_t. For a first order phase transition, since there is a discontinuity of the first derivative of ϕ at β_t, the specific heat is not well defined and one speaks rather of latent heat, related to the jump $(\varepsilon_2 - \varepsilon_1)$ of the energy as shown in Fig. 8b. At second order phase transitions, the second derivative of ϕ is instead well defined and is discontinuous at β_t.

As explained above, the presence of a convex "intruder" like in Fig. 8a in the entropy of long-range systems does not imply the appearance of singularities in the entropy, and therefore it could be doubted that this behavior signals a true phase transition in the microcanonical ensemble. Since this feature has been found first for gravitational systems, it is sometimes called "gravitational phase transition". We will better clarify this issue analyzing what happens in the canonical ensemble. The rescaled free energy $\phi(\beta)$ is again expressed by Eq. (57) and the mean value of the energy is

$$\varepsilon(\beta) = \frac{\partial \phi}{\partial \beta}. \tag{72}$$

The plot of $\varepsilon(\beta)$ is obtained from the curve in Fig. 8b by considering the ordinate β as the control variable. The Maxwell construction is realized by the horizontal dashed line at $\beta(\varepsilon_1)$. If, in the canonical ensemble, we start from a value of β such that the energy of the system is less than ε_1 and we gradually decrease β, the system will reach the energy ε_1 and then will jump to the energy ε_2, and after the jump β will decrease continuously. Therefore, while in the microcanonical ensemble there is no singularity of the entropy, in the canonical ensemble there is a discontinuity of the derivative of the rescaled free energy ϕ, corresponding to a jump in the energy (associated to a latent heat). In the canonical ensemble the system has therefore a first order phase transition. Equilibrium macroscopic states with energies in the range $[\varepsilon_1, \varepsilon_2]$ do not exist, since the lack of additivity, as we noted, does not allow, contrary to short-range systems, to have mixtures of states as in Eq. (63). The temperature of the phase transition in the canonical ensemble is obtained by the Maxwell construction. Correspondingly, microcanonical microstates exist in the energy range $[\varepsilon_1, \varepsilon_2]$ and *phase separation* is not thermodynamically favored in this ensemble.

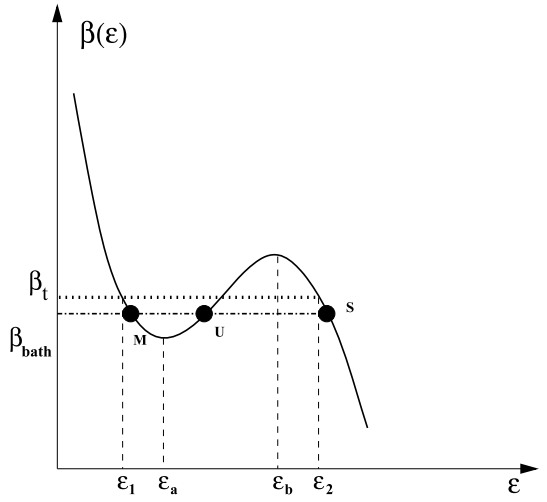


Fig. 9. Inverse temperature β as a function of energy ε. The Maxwell's construction is shown by the dotted line, while the dash–dotted line indicates the inverse temperature of the bath β_{bath}. U denotes an unstable macroscopic state with negative specific heat, while M and S are metastable and stable, respectively.

The fact that the presence of a canonical first order phase transition is necessary to obtain ensemble inequivalence was conjectured in Ref. [14]. This statement has been put on a more rigorous basis in Refs. [20,42], analyzing the convexity properties of the entropy $s(\varepsilon)$. In fact, it has been shown [20] that if the rescaled free energy $\phi(\beta)$ is differentiable, then the entropy $s(\varepsilon)$ can be obtained by its Legendre–Fenchel transform. This applies also for second order phase transitions, when the second derivative of $\phi(\beta)$ is discontinuous. Therefore, in the presence of a second order phase transition in the canonical ensemble, the microcanonical and canonical ensembles are equivalent.

In this subsection we have discussed in detail the case where no singularity is present in the entropy. Although already showing all the features of ensemble inequivalence, this case is not generic and we will discuss in the next subsection a model that has a second order phase transition in the microcanonical ensemble and still a first order transition in the canonical ensemble.

Let us conclude this subsection with a remark. We have remarked that energies between ε_1 and ε_2 correspond to the same value of β in the canonical ensemble. It is interesting to figure out what happens if an initially isolated system with negative specific heat and with an energy between ε_1 and ε_2, is put in contact with a heat bath that has its inverse temperature β_{bath}. Looking at Fig. 9 can be of help to understand the argument. We consider the case where the energy of the system lies in the range in which the specific heat is negative $[\varepsilon_a, \varepsilon_b]$ when the system is put in contact with the bath. Let us take for instance point U in Fig. 9 as an initial point. We are interested to study the behavior of the system subjected to small perturbations, so that it can still be considered to be initially close to a microcanonical system. We see immediately that the system becomes unstable. In fact, if it gets a small amount of energy from the bath, its temperature lowers (negative specific heat!), and therefore further energy will flow from the bath to the system, inducing a lowering of the system's temperature and then creating instability. If, on the contrary, the initial energy fluctuation decreases the system's energy, its temperature rises, inducing a further energy flow towards the bath, and, hence, an increase of system's temperature. Thus, in contact with a heat bath, the system does not maintain an energy in which its microcanonical specific heat is negative. The flow of energy started by the initial energy fluctuations stops when the system reaches the same temperature of the bath again, but at an energy for which its specific heat is positive. Looking at Fig. 9, it is clear that this could be either outside the range $[\varepsilon_1, \varepsilon_2]$, i.e. point S, or inside this range, point M. This feature is valid for all points U inside $]\varepsilon_a, \varepsilon_b[$. Once in M, the system will be in a thermodynamically metastable state and a sufficiently large fluctuation in the energy exchange with the bath will make it leave this metastable state, ending up again in a state with energy outside $[\varepsilon_1, \varepsilon_2]$, i.e. point S, which has the same inverse temperature of the bath β_{bath}. If the system instead jumps directly from U to S, it will stay there because this point lies on a thermodynamically stable branch.

4.2. An analytical solvable example: The mean-field Blume–Emery–Griffiths model

We have presented above the main physical and mathematical aspects related to ensemble equivalence or inequivalence in the study of long-range systems. Other mathematical approaches and tools, that exist, will be presented in connection with concrete examples. Actually, this subsection is dedicated to a toy model that exhibits all the features that have been discussed so far, in particular ensemble inequivalence and negative specific heat in the microcanonical ensemble. Historically, the relation between first order phase transition and negative specific heat for long-range systems in the thermodynamic limit was first pointed out in Refs. [104,105]. The phenomenology we are going to discuss in this section has been heuristically described in Ref. [106].

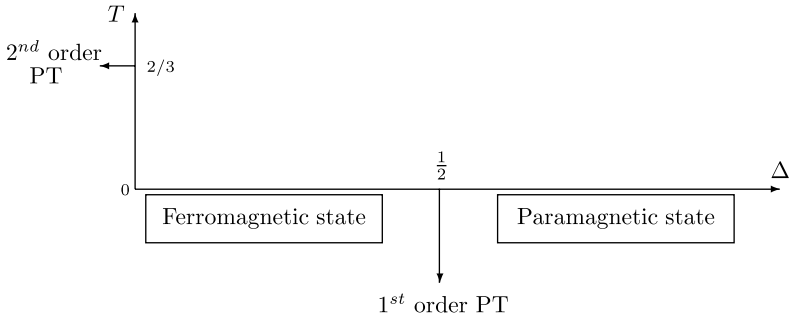


Fig. 10. Elementary features of the phase diagram of the Blume–Emery–Griffiths model, showing the phase transitions on the temperature T and local coupling Δ axis, respectively.

4.2.1. Qualitative remarks

The Blume–Emery–Griffiths (BEG) model is a lattice spin model with infinite range, mean-field like interactions whose phase diagram can be obtained analytically both within the canonical and the microcanonical ensembles. This study enables one to compare the two resulting phase diagrams and get a better understanding of the effect of the non-additivity on the thermodynamic behavior of the model.

The model we consider is a simplified version of the Blume–Emery–Griffiths model [107], known as the Blume–Capel model, where the quadrupole–quadrupole interaction is absent. The model is intended to reproduce the relevant features of superfluidity in He^3–He^4 mixtures. Recently, it has also been proposed as a realistic model for metallic ferromagnetism [108]. It is a lattice system (5), and each lattice point i is occupied by a spin-1 variable, i.e., a variable S_i assuming the values $S_i = 0, \pm 1$. We will consider the mean-field version of this model, for which all lattice points are coupled with the same strength. The Hamiltonian is given by

$$H = \Delta \sum_{i=1}^{N} S_i^2 - \frac{J}{2N} \left(\sum_{i=1}^{N} S_i \right)^2 , \tag{73}$$

where $J > 0$ is a ferromagnetic coupling constant and $\Delta > 0$ controls the energy difference between the ferromagnetic $S_i = 1, \forall i$, or $S_i = -1, \forall i$, and the paramagnetic, $S_i = 0, \forall i$, states. In the following we will set $J = 1$, without loss of generality since we consider only ferromagnetic couplings. The paramagnetic configuration has zero energy, while the uniform ferromagnetic configurations have an energy $(\Delta - 1/2)N$. In the canonical ensemble, the minimization of the free energy $F = E - TS$ at zero temperature is equivalent to the minimization of the energy. One thus finds that the paramagnetic state is the more favorable from the thermodynamic point of view if $E(\{\pm 1\}) > E(\{0\})$, which corresponds to $\Delta > 1/2$. At the point $\Delta = 1/2$, there is therefore a phase transition; it is a *first* order phase transition since, it corresponds to a sudden jump of magnetization from the ferromagnetic state to the paramagnetic state.

For vanishingly small ratio Δ, the first term of Hamiltonian (73) can be safely neglected so that one recovers the Curie–Weiss Hamiltonian (1) with spin 1, usually introduced to solve the Ising model within the mean-field approximation. It is well known that such a system has a *second* order phase transition when $T = 2/3$ (we recall that we are adopting units for which $J = 1$, $k_B = 1$). Since one has phase transitions of different orders on the T and Δ axis (see Fig. 10), one expects that the (T, Δ) phase diagram displays a *transition line* separating the low temperature ferromagnetic phase from the high temperature paramagnetic phase. The transition line is indeed found to be first order at large Δ values, while it is second order at small Δ's.

4.2.2. The solution in the canonical ensemble

The canonical phase diagram of this model in the (T, Δ) is known since long time [107,109,110]. The partition function reads

$$Z(\beta, N) = \sum_{\{S_1, \ldots, S_N\}} \exp \left(-\beta \Delta \sum_{i=1}^{N} S_i^2 + \frac{\beta J}{2N} \left(\sum_{i=1}^{N} S_i \right)^2 \right). \tag{74}$$

Using the Gaussian identity

$$\exp(bm^2) = \sqrt{\frac{b}{\pi}} \int_{-\infty}^{+\infty} \mathrm{d}x \exp(-bx^2 + 2mbx), \tag{75}$$

(often called the Hubbard–Stratonovich transformation) with $m = \sum_i S_i/N$ and $b = N\beta J/2$, one obtains

$$Z(\beta, N) = \sum_{\{S_1, \ldots, S_N\}} \exp \left(-\beta \Delta \sum_{i=1}^{N} S_i^2 \right) \sqrt{\frac{N\beta}{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}x \exp \left(-\frac{N\beta}{2} x^2 + mN\beta x \right). \tag{76}$$

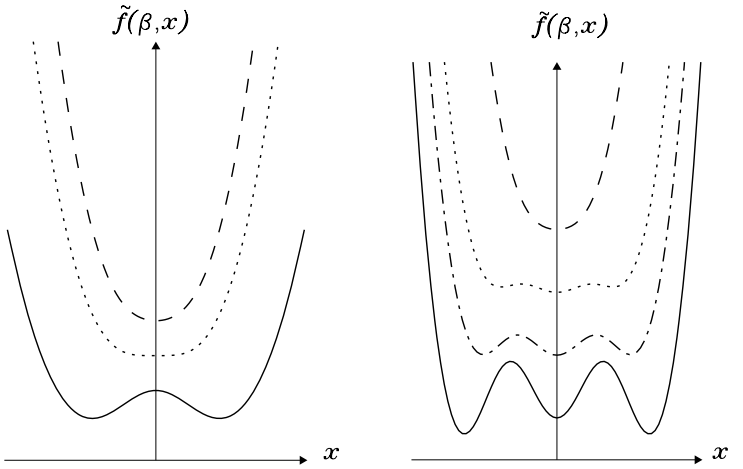


Fig. 11. Free energy $\tilde{f}(\beta, x)$ vs. x for different values of the inverse temperature $\beta = 1/T$. Left panel shows the case of a *second* order phase transition, temperature values $T = 0.8$ (dashed line), 0.63 (dotted), 0.4 (solid) when $\Delta = 0.1$ are displayed. Right panel shows the case of a *first* order phase transition with $\Delta = 0.485$ when $T = 0.5$ (dashed), 0.24 (dotted), 0.21 (dash–dotted), 0.18 (solid).

One then easily gets

$$Z(\beta, N) = \sqrt{\frac{N\beta}{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}x \exp(-N\beta\tilde{f}(\beta, x)) \tag{77}$$

where

$$\tilde{f}(\beta, x) = \frac{1}{2}x^2 - \frac{1}{\beta}\ln[1 + \mathrm{e}^{-\beta\Delta}(\mathrm{e}^{\beta x} + \mathrm{e}^{-\beta x})]. \tag{78}$$

The integral in (77) can be computed using the saddle point method where N is the large parameter. The free energy is thus

$$f(\beta) = \inf_x \tilde{f}(\beta, x). \tag{79}$$

It is not difficult to see that the spontaneous magnetization $\langle m\rangle$ is equal to the value of x at the extremum which appears in Eq. (79). We should also note that $\tilde{f}(\beta, x)$ is even in x; therefore, if there is a value of x different from 0 realizing the extremum, the opposite value also realizes it. This means that if the minimum $\bar{x}$ is equal to 0 the system is in the paramagnetic phase, while if $\bar{x} \neq 0$ the system is in the ferromagnetic phase, where it can assume a positive or a negative magnetization. The phase diagram, in the (T, Δ) plane, is then divided into a paramagnetic region ($\bar{x} = 0$) and a ferromagnetic one ($\bar{x} \neq 0$).

Let us now show that the two regions are divided by a second order phase transition line and a first order phase transition line, which meet at a tricritical point. As in the Landau theory of phase transitions, we find a second order transition line by a power series expansion in x of the function $\tilde{f}(\beta, x)$ in Eq. (78). The second order line is obtained by equating to zero the coefficient of x^2, i.e., by the relation

$$A_c \equiv \beta - \frac{1}{2}\mathrm{e}^{\beta\Delta} - 1 = 0, \tag{80}$$

provided that the coefficient of x^4 is positive, i.e., provided that

$$B_c \equiv 4 - \mathrm{e}^{\beta\Delta} > 0. \tag{81}$$

The tricritical point is obtained when $A_c = B_c = 0$. This gives $\Delta = \ln(4)/3 \simeq 0.4621$ and $\beta = 3$. The continuation of the critical line after the tricritical point is the first order phase transition line, which can be obtained by finding numerically the local maximum value $\bar{x} \neq 0$ (magnetic phase) for which $\tilde{f}(\beta, x)$ is equal to $\tilde{f}(\beta, 0)$ (paramagnetic phase), i.e., by equating the free energies of the ferromagnetic and the paramagnetic phases. The behavior of the function $\tilde{f}(\beta, x)$ as β varies is shown in Fig. 11: panel (a) represents the case of a second order phase transition ($\Delta = 0.1$) and panel (b) the case of a first order phase transition ($\Delta = 0.485$).

A picture of the phase diagram is shown in Fig. 12.

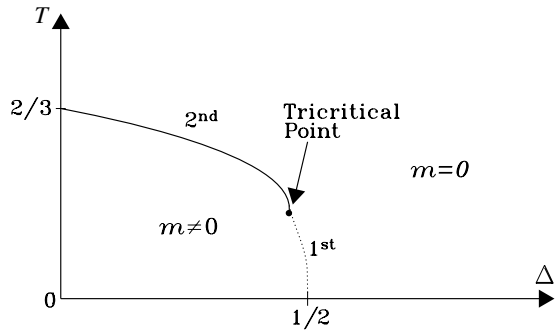


Fig. 12. Phase diagram of the Blume–Emery–Griffiths model in the canonical ensemble. The second order transition line (solid) ends at the tricritical point (•), where the transition becomes first order (dotted).

4.2.3. The solution in the microcanonical ensemble

The derivation of the phase diagram of the BEG model (73) in the *microcanonical ensemble* relies on a simple counting problem [14], since all spins interact with equal strength, independently of their mutual distance. A given macroscopic configuration is characterized by the numbers N_+, N_-, N_0 of up, down and zero spins, with $N_+ + N_- + N_0 = N$. The energy E of this configuration is only a function of N_+, N_- and N_0 and is given by

$$E = \Delta Q - \frac{1}{2N} M^2, \tag{82}$$

where $Q = \sum_{i=1}^{N} S_i^2 = N_+ + N_-$ (the quadrupole moment) and $M = \sum_{i=1}^{N} S_i = N_+ - N_-$ (the magnetization) are the two order parameters. The number of microscopic configurations Ω compatible with the macroscopic occupation numbers N_+, N_- and N_0 is

$$\Omega = \frac{N!}{N_+! N_-! N_0!}. \tag{83}$$

Using Stirling's approximation in the large N limit, the entropy, $S = \ln \Omega$, is given by

$$S = -N \left[(1-q)\ln(1-q) + \frac{1}{2}(q+m)\ln(q+m) + \frac{1}{2}(q-m)\ln(q-m) - q\ln 2 \right], \tag{84}$$

where $q = Q/N$ and $m = M/N$ are the quadrupole moment and the magnetization per site, respectively. Eq. (82) may be written as

$$q = 2K\varepsilon + Km^2, \tag{85}$$

where $K = 1/(2\Delta)$. Using this relation, the entropy per site $\tilde{s} = S/N$ can be expressed in terms of m and ε, as follows

$$\begin{aligned} \tilde{s}(\varepsilon, m) = & -(1 - 2K\varepsilon - Km^2)\ln(1 - 2K\varepsilon - Km^2) - \frac{1}{2}(2K\varepsilon + Km^2 + m)\ln(2K\varepsilon + Km^2 + m) \\ & - \frac{1}{2}(2K\varepsilon + Km^2 - m)\ln(2K\varepsilon + Km^2 - m) + (2K\varepsilon + Km^2)\ln 2. \end{aligned} \tag{86}$$

At fixed ε, the value of m which maximizes the entropy corresponds to the equilibrium magnetization. The corresponding equilibrium entropy

$$s(\varepsilon) = \sup_m \tilde{s}(\varepsilon, m) \tag{87}$$

contains all the relevant information about the thermodynamics of the system in the microcanonical ensemble. As usual in systems where the energy per particle is bounded from above, the model has both a positive and a negative temperature region: entropy is a one humped function of the energy. In order to locate the continuous transition line, one develops $\tilde{s}(\varepsilon, m)$ in powers of m, in analogy with what has been done above for the canonical free energy

$$\tilde{s} = \tilde{s}_0 + A_{mc}\, m^2 + B_{mc}\, m^4 + O(m^6), \tag{88}$$

where

$$\tilde{s}_0 = \tilde{s}(\varepsilon, m = 0) = -(1 - 2K\varepsilon)\ln(1 - 2K\varepsilon) - 2K\varepsilon \ln(K\varepsilon), \tag{89}$$

and

$$A_{mc} = -K \ln \frac{K\varepsilon}{(1-2K\varepsilon)} - \frac{1}{4K\varepsilon}, \tag{90}$$

$$B_{mc} = -\frac{K}{4\varepsilon(1-2K\varepsilon)} + \frac{1}{8K\varepsilon^2} - \frac{1}{96K^3\varepsilon^3}. \tag{91}$$

In the paramagnetic phase both A_{mc} and B_{mc} are negative, and the entropy is maximized by $m = 0$. The continuous transition to the ferromagnetic phase takes place at $A_{mc} = 0$ for $B_{mc} < 0$. In order to obtain the critical line in the (T, Δ) plane, we first observe that temperature is calculable on the critical line ($m = 0$) using (53) and (89). One gets

$$\frac{1}{T} = 2K \ln \frac{1-2K\varepsilon}{K\varepsilon}. \tag{92}$$

Requiring now that $A_{mc} = 0$, one gets the following expression for the critical line

$$\beta = \frac{\exp[\beta/(2K)]}{2} + 1. \tag{93}$$

Equivalently, this expression may be written as $\beta = 1/(2K\varepsilon)$. The microcanonical critical line thus coincides with the critical line (80) obtained for the canonical ensemble. The tricritical point of the microcanonical ensemble is obtained at $A_{mc} = B_{mc} = 0$. Combining these equations with Eq. (92), one finds that, at the tricritical point, β satisfies the equation

$$\frac{K^2}{2\beta^2}\left[1 + 2\exp\left(-\frac{\beta}{2K}\right)\right] - \frac{K}{2\beta} + \frac{1}{12} = 0. \tag{94}$$

Eqs. (93) and (94) yield a tricritical point at $K \simeq 1.0813$, $\beta = 3.0272$. This has to be compared with the canonical tricritical point located at $K = 1/(2\Delta) = 3/\ln(16) \simeq 1.0820$, $\beta = 3$. The two points, although very close to each other, do not coincide. The microcanonical critical line extends beyond the canonical one. This feature, which is a clear indication of ensemble inequivalence, was first found analytically for the BEG model [14] and later confirmed for gravitational models [4,34]. The non-coincidence of microcanonical and canonical tricritical points is a generic feature, as proven in Ref. [42].

4.2.4. Inequivalence of ensembles

We have already discussed in general terms the question of ensemble equivalence or inequivalence in Sections 4.1.1 and 4.1.3. Inequivalence is associated to the existence of a convex region of the entropy as a function of energy. This is exactly what happens for the BEG model in the region of parameters $1 < K < 3\ln(16)$. Since the interesting region is extremely narrow for this model [14], it is more convenient to plot a schematic representation of the entropy and of the free energy (see Fig. 13). We show what happens in a region of K where both a *negative specific heat* and a *temperature jump* are present. The entropy curve consists of two branches: the high energy branch is obtained for $m = 0$ (dotted line), while the low energy one is for $m \neq 0$ (full line). The $m = 0$ branch has been extended also in a region where it corresponds to metastable states, just to emphasize that these correspond to a smaller entropy and that it remains a concave function overall the energy range. We have not extended the $m \neq 0$ branch in the high energy region not to make the plot confusing: it would also correspond to a metastable state. The two branches merge at an energy value ε_t where the left and right derivatives do not coincide; hence microcanonical temperature is different on the two sides, leading to a *temperature jump*. It has been proven in Ref. [42], that for all types of bifurcation the temperature jump is always negative. In the low energy branch, there is a region where entropy is locally convex (thick line in Fig. 13), giving a *negative specific heat* according to formula (70). The convex envelope, with constant slope β_t is also indicated by the dash–dotted line. In the same figure, we plot the rescaled free energy $\phi(\beta)$, which is a concave function, with a point β_t where left and right derivatives (given by ε_1 and ε_2 respectively) are different. This is the first order phase transition point in the canonical ensemble.

A schematic phase diagram near the canonical tricritical point (CTP) and the microcanonical one (MTP) is given in Fig. 14. In the region between the two tricritical points, the canonical ensemble yields a first order phase transition at a higher temperature, while in the microcanonical ensemble the transition is still continuous. It is in this region that negative specific heat appears. Beyond the microcanonical tricritical point, temperature has a jump at the transition energy in the microcanonical ensemble. The two lines emerging on the right side from the MTP correspond to the two limiting temperatures which are reached when approaching the transition energy from below and from above (see Fig. 15c and d). The two microcanonical temperature lines and the canonical first order phase transition line all merge on the $T = 0$ line at $\Delta = 1/2$.

To get a better understanding of the microcanonical phase diagram and also in order to compare our results with those obtained for self-gravitating systems [4,34] and for finite systems [9,19,31,87], we consider the temperature–energy relation $T(\varepsilon)$ (also called in the literature the "caloric curve"). Also this curve has two branches: a high energy branch (92) corresponding to $m = 0$, and a low energy branch obtained from (53) using the spontaneous magnetization $m_s(\varepsilon) \neq 0$. At the intersection point of the two branches, the two entropies become equal. However, their first derivatives at the crossing point can be different, resulting in a jump in the temperature, i.e. *a microcanonical first order transition*. When the transition is

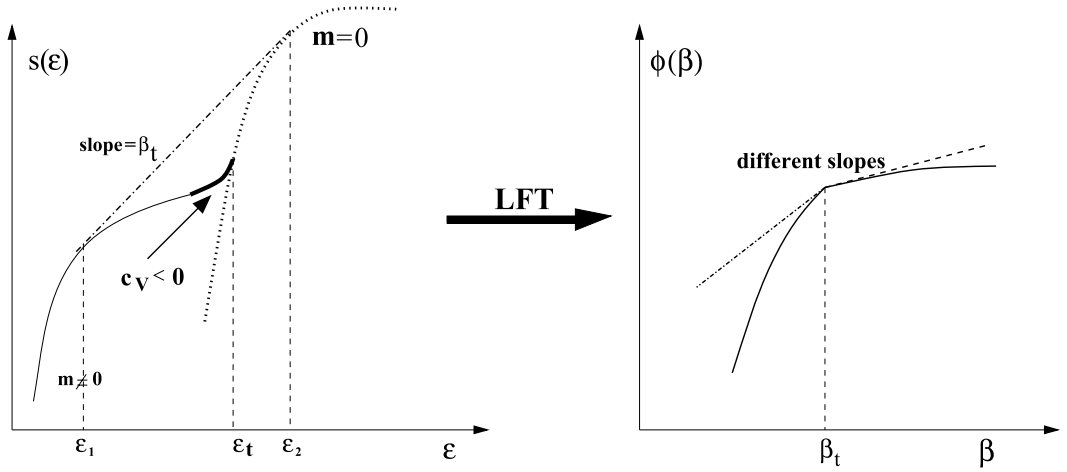


Fig. 13. Left graph: schematic plot of the entropy $s(\varepsilon)$ as a function of energy density ε for the BEG model in a case where negative specific heat coexists with a temperature jump. The dash–dotted line is the concave envelope of $s(\varepsilon)$ and the region with negative specific heat $c_V < 0$ is explicitly indicated by the thick line. Right graph: Rescaled free energy $\phi(\beta)$: the first order phase transition point β_t is shown.

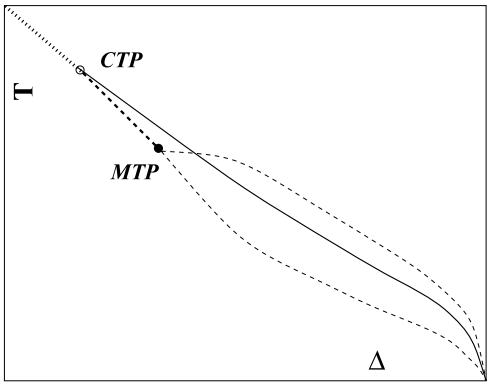


Fig. 14. A schematic representation of the phase diagram, where we expand the region around the canonical (CTP) and the microcanonical (MTP) tricritical points. The second order line, common to both ensembles, is dotted, the first order canonical transition line is solid and the microcanonical transition lines are dashed (with the bold dashed line representing a continuous transition).

continuous in the microcanonical ensemble, i.e. the first derivative of the entropy branches at the crossing point are equal, the BEG model always displays, at variance with what happens for gravitational systems, a discontinuity in the second derivative of the entropy. This is due to the fact that here we have a true symmetry breaking transition [42]. Fig. 15 displays the $T(\varepsilon)$ curves for decreasing values of K. For $K = 3/\ln(16)$, corresponding to the canonical tricritical point, the lower branch of the curve has a zero slope at the intersection point (Fig. 15a). Thus, the specific heat of the ordered phase diverges at this point. This effect signals the canonical tricritical point through a property of the microcanonical ensemble. Decreasing K, down to the region between the two tricritical points, a *negative specific heat* in the microcanonical ensemble first arises ($\partial T/\partial\varepsilon < 0$), see Fig. 15b. At the microcanonical tricritical point, the derivative $\partial T/\partial\varepsilon$ of the lower branch diverges at the transition point, yielding a vanishing specific heat. For smaller values of K, a jump in the temperature appears at the transition energy (Fig. 15c). The lower temperature corresponds to the $m = 0$ solution (92) and the upper one is given by $\exp(\beta/2K) = 2(1-q^*)/\sqrt{(q^*)^2 - (m^*)^2}$, where m^*, q^* are the values of the order parameters of the ferromagnetic state at the transition energy. The negative specific heat branch disappears at even smaller values of K, leaving just a temperature jump (see Fig. 15d). In the $K \to 1$ limit the low temperature branch, corresponding to $q = m = 1$ in the limit, shrinks to zero and the $m = 0$ branch (92) occupies the full energy range.

4.3. Entropy and free energy dependence on the order parameter

In this section, we will discuss in detail the dependence of both the canonical free energy and the microcanonical entropy on the order parameter. This will allow to understand more deeply the relation between the two ensembles by revisiting Maxwell constructions. Besides that, we will also discover an interesting physical effect, *negative susceptibility*, of which we will give an explicit example, see Section 4.5.4.

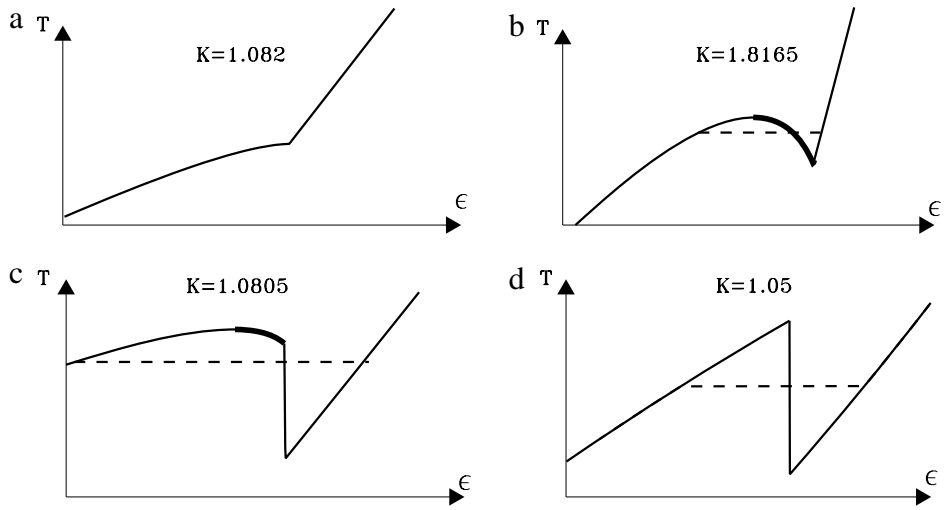


Fig. 15. The temperature–energy relation in the microcanonical ensemble for different values of K. The dashed horizontal line is the Maxwell construction in the canonical ensemble and identifies the canonical first order transition temperature at the point where two minima of the free energy coexist. Thick lines identify negative specific heat in the microcanonical ensemble. We do not report the numerical values on the axes for the readability of the figure.

4.3.1. Basic definitions

Let us start from Eq. (86), that gives the entropy per site $\tilde{s}(\varepsilon, m)$, for the BEG model, as a function of the energy per site ε *and* the magnetization m. This entropy is proportional to the logarithm of the number of configurations which have a given energy *and* a given magnetization. In the general expression Eq. (50), these configurations can be obtained by adding a further Dirac delta function in the integrand, so that only the configurations with a given m would be counted. Thus one gets

$$\tilde{s}(\varepsilon, m) = \lim_{N\to\infty} \frac{1}{N} \ln \int \prod_i^N \mathrm{d}\mathbf{q}_i \delta\left(E - U(\{\mathbf{q}_i\})\right) \delta\left(Nm - M(\{\mathbf{q}_i\})\right) \tag{95}$$

where M is the total magnetization corresponding to configuration $\{\mathbf{q}_i\}$. For spin models, the local variable takes discrete values $\mathbf{q}_i \equiv S_i$, hence $M = \sum_i S_i$ and the integral in Eq. (95) is replaced by a discrete sum. Besides that, there is no kinetic energy, hence only the potential energy appears $H = U$. The calculation of entropy (95) is often an intermediate step in the calculation of $s(\varepsilon)$, with M the order parameter. In order to get $s(\varepsilon)$, one computes the global maximum of the constrained entropy (95). In the thermodynamic limit, this procedure is fully justified, since the relative contribution of all configurations corresponding to values of the order parameter that are different from the one realizing the global maximum, vanishes. This is what has been done in our study of the BEG model, e.g. in Eq. (87).

In the canonical ensemble, the computation of the partition function for a given value of the order parameter, i.e. for the system at a given temperature *and* at a given magnetization m, can be obtained by adding a Dirac delta function to the integrand in Eq. (55)

$$\tilde{f}(\beta, m) = -\frac{1}{\beta} \lim_{N\to\infty} \frac{1}{N} \ln \sum_{\{S_1,\ldots,S_N\}} \mathrm{d}S_i \exp\left[-\beta H(\{S_i\})\right] \delta\left(Nm - M(\{S_i\})\right). \tag{96}$$

Thus, the free energy depends on both β *and* the magnetization m. Finally, a relation analogous to Eq. (57) holds between the entropy and free energy

$$\tilde{f}(\beta, m) = \inf_{\varepsilon} \left[\varepsilon - \frac{1}{\beta}\tilde{s}(\varepsilon, m)\right], \tag{97}$$

valid, as before, for all systems, independently of the range of the interactions.

We are therefore led to the introduction of the *generalized free energy* (see Eq. (68))

$$\widehat{f}(\beta, \varepsilon, m) = \varepsilon - \frac{1}{\beta}\tilde{s}(\varepsilon, m), \tag{98}$$

that will be used, as previously, to study the relation between microcanonical and canonical equilibrium states. Needless to say, it is not at all guaranteed that the function $\tilde{s}(\varepsilon, m)$ can be easily derived, in general, as we have done for the BEG model; nevertheless this general discussion is useful to show how the properties of this function explain the occurrence, or not, of ensemble equivalence.

In the microcanonical ensemble, the entropy $s(\varepsilon)$ of the system at a given energy ε is given by formula (87). In the canonical ensemble, the free energy $f(\beta)$ of the system at a given inverse temperature β will be given by

$$f(\beta) = \inf_{\varepsilon,m} \widehat{f}(\beta, \varepsilon, m) = \inf_{m} \tilde{f}(\beta, m) = \inf_{\varepsilon,m} \left[\varepsilon - \frac{1}{\beta} \tilde{s}(\varepsilon, m) \right], \tag{99}$$

as can be easily deduced by Eqs. (87), (57), (97) and (98). The two extremal problems (87) and (99), that basically contain the single function $\tilde{s}(\varepsilon, m)$, can be employed to study ensemble equivalence. Suppose we fix β and solve the extremal problem (99), finding the values of ε and m that realize an extremum. Then, we will have ensemble equivalence if the same values will realize the extremum in formula (87) while at the same time the derivative $\partial s/\partial \varepsilon$ will be equal to the fixed value of β. In conclusion, we seek in both extremal problems the solution of the following first order conditions

$$\frac{\partial \tilde{s}}{\partial m} = 0 \tag{100}$$

$$\frac{\partial \tilde{s}}{\partial \varepsilon} = \beta. \tag{101}$$

We denote by $\varepsilon^*(\beta)$, $m^*(\beta)$ the solution of the variational problem (100) and (101). Using (101), it is straightforward to verify that

$$\frac{\mathrm{d}(\beta f)}{\mathrm{d}\beta} = \varepsilon^*(\beta), \tag{102}$$

meaning that the value of ε at the extremum is indeed the canonical mean energy.

However, we have to consider also the stability of these extrema. We denote derivatives by subscripts, e.g., $\tilde{s}_m$ is the first derivative of $\tilde{s}$ with respect to m. The only condition required by (87) is that $\tilde{s}_{mm} < 0$. In order to discuss the stability of the canonical solution, one has to determine the sign of the eigenvalues of the Hessian of the function to be minimized in (99). The Hessian is

$$\mathcal{H} = -\frac{1}{\beta} \begin{pmatrix} \tilde{s}_{mm} & \tilde{s}_{m\varepsilon} \\ \tilde{s}_{\varepsilon m} & \tilde{s}_{\varepsilon\varepsilon} \end{pmatrix}. \tag{103}$$

The extremum is a minimum if and only if both the determinant and the trace of the Hessian are positive

$$-\tilde{s}_{\varepsilon\varepsilon} - \tilde{s}_{mm} > 0 \tag{104}$$

$$\tilde{s}_{\varepsilon\varepsilon}\tilde{s}_{mm} - \tilde{s}_{m\varepsilon}^2 > 0. \tag{105}$$

This implies that $\tilde{s}_{\varepsilon\varepsilon}$ and $\tilde{s}_{mm}$ must be negative, and moreover $\tilde{s}_{\varepsilon\varepsilon} < -\tilde{s}_{m\varepsilon}^2/|\tilde{s}_{mm}|$. This has strong implications on the canonical specific heat, which must be positive, as it has been shown on general grounds in Section 4.1.3, see Eq. (19). Let us prove it by using the variational approach, instead of the usual Thirring argument [11,12], that uses the expression of the canonical partition sum. Indeed, taking the derivatives of Eqs. (100) and (101) with respect to β, after having substituted into them $\varepsilon^*(\beta)$, $m^*(\beta)$, one gets

$$\tilde{s}_{\varepsilon\varepsilon}\frac{\mathrm{d}\varepsilon^*}{\mathrm{d}\beta} + \tilde{s}_{\varepsilon m}\frac{\mathrm{d}m^*}{\mathrm{d}\beta} = 1 \tag{106}$$

$$\tilde{s}_{m\varepsilon}\frac{\mathrm{d}\varepsilon^*}{\mathrm{d}\beta} + \tilde{s}_{mm}\frac{\mathrm{d}m^*}{\mathrm{d}\beta} = 0, \tag{107}$$

where all second derivatives are computed at $\varepsilon^*(\beta)$, $m^*(\beta)$. Recalling now that the specific heat per particle at constant volume is

$$c_V = \frac{\mathrm{d}\varepsilon^*}{\mathrm{d}T} = -\beta^2 \frac{\mathrm{d}\varepsilon^*}{\mathrm{d}\beta}, \tag{108}$$

one gets

$$c_V = \beta^2 \frac{\tilde{s}_{mm}}{\tilde{s}_{\varepsilon m}^2 - \tilde{s}_{\varepsilon\varepsilon}\tilde{s}_{mm}}, \tag{109}$$

which is always positive if the stability conditions (104) and (105) are satisfied. Since the stability condition in the microcanonical ensemble only requires that $\tilde{s}_{mm} < 0$, a canonically stable solution is also microcanonically stable. The converse is not true: one may well have an entropy maximum, $\tilde{s}_{mm} < 0$, which is a free energy saddle point, with $\tilde{s}_{\varepsilon\varepsilon} > 0$. This implies that the specific heat (109) can be negative.

The above results are actually quite general, provided the canonical and microcanonical solutions are expressed through variational problems of the type (87) and (99). The extrema, and thus the caloric curves $T(\varepsilon)$, are the same in the two ensembles, but the stability of the different branches is different. This aspect was first discussed by Katz [111] in connection with self-gravitating systems (see also Ref. [34]).

4.3.2. Maxwell construction in the microcanonical ensemble

We have already discussed Maxwell's construction in Section 4.1.2. We have shown that, for short range systems, where microcanonical and canonical ensembles are always equivalent, Maxwell's construction derives from the concave envelope construction for the microcanonical entropy. This in turn is a consequence of *additivity* and of the presence of a first order phase transition in the canonical ensemble. Here, we discuss Maxwell's construction for long-range systems, where the microcanonical entropy can have a stable convex intruder, leading to ensemble inequivalence.

The study of the BEG model, in Section 4.2, has emphasized the presence of an extremely rich phenomenology. In particular, in a specific region of the control parameter K, the canonical and microcanonical ensembles show a first order phase transition, with a forbidden energy range in the former ensemble and a temperature jump in the latter (see Fig. 15c). Both the $\varepsilon(\beta)$ curve and the $\beta(\varepsilon)$ one become multiply valued if we include metastable and unstable states. Since we know that the Maxwell construction leads to an equal area condition for the $\beta(\varepsilon)$ curve, which defines the phase transition inverse temperature β_t in the canonical ensemble, we wonder here whether a similar construction exists for the $\varepsilon(\beta)$ relation which would lead to the determination of the transition energy ε_t in the microcanonical ensemble.

In the following discussion of Maxwell's construction, it is crucial to understand the mechanism that generates multiple branches of the $\beta(\varepsilon)$ curve. Let us define

$$\tilde{\beta}(\varepsilon, m) = \frac{\partial \tilde{s}(\varepsilon, m)}{\partial \varepsilon}. \tag{110}$$

We have explained that the equilibrium magnetization m^*, at any energy ε, is the global maximum of the entropy per site (86). Once m^* has been computed, the equilibrium inverse temperature is given by $\beta(\varepsilon) = \tilde{\beta}(\varepsilon, m^*)$. However, local maxima, local minima and saddles of the entropy also exist, corresponding to different values of m. Following such critical points as a function of ε, one determines the different branches of $\beta(\varepsilon)$. In particular, we have the continuation at energies lower than the microcanonical transition energy ε_t of the high energy $m = 0$ branch (dashed part of $\beta_H(\varepsilon)$ in Fig. 16) and the continuation at higher energies of the magnetized branch $\beta_L(\varepsilon)$. It is interesting to remark that the $m = 0$ point remains an extremum for all values of ε since $\tilde{s}(\varepsilon, m)$ is even in m.

An example of inverse temperature $\beta(\varepsilon)$ relation is plotted in Fig. 16. The lower branch β_L starts at low energy and ends at the energy ε_H, where its derivative becomes infinite. The upper branch β_H starts at high energy and ends at the energy ε_L, where again its derivative is infinite. These two branches are connected by the vertical line at energy ε_t, and by the intermediate branch β_I that goes from ε_L to ε_H. The equilibrium state is given by the lower branch for $\varepsilon < \varepsilon_t$ and by the upper branch for $\varepsilon > \varepsilon_t$. The lower branch for $\varepsilon_t < \varepsilon < \varepsilon_H$ (dashed) and the upper branch for $\varepsilon_L < \varepsilon < \varepsilon_c$ (dashed) represent metastable states, while the intermediate branch represents unstable states (dotted). Therefore, increasing the energy, the equilibrium value of β jumps from the lower to the upper branch (thus following the vertical line) at the transition energy ε_t. It is easy to show that the vertical line realizes a Maxwell construction, i.e., that the two areas A_1 and A_2 are equal. The curve $\beta(\varepsilon)$ has therefore three branches, that we denote by $\beta_L(\varepsilon)$ (the low energy magnetized branch), $\beta_I(\varepsilon)$ (the intermediate branch of unstable states) and $\beta_H(\varepsilon)$ (the high energy paramagnetic branch). Then, we have

$$A_2 - A_1 = \int_{\varepsilon_t}^{\varepsilon_H} \beta_L(\varepsilon)\mathrm{d}\varepsilon + \int_{\varepsilon_H}^{\varepsilon_L} \beta_I(\varepsilon)\mathrm{d}\varepsilon + \int_{\varepsilon_L}^{\varepsilon_t} \beta_H(\varepsilon)\mathrm{d}\varepsilon \tag{111}$$

$$= (s_L(\varepsilon_H) - s_L(\varepsilon_t)) + (s_I(\varepsilon_L) - s_I(\varepsilon_H)) + (s_H(\varepsilon_t) - s_H(\varepsilon_L)), \tag{112}$$

where in the r.h.s. $s_i(\varepsilon)$ is the function whose derivative gives the branch $\beta_i(\varepsilon)$, with $i = H, I, L$. We use now the continuity property of the entropy, imposing that $s_L(\varepsilon_H) = s_I(\varepsilon_H)$ and $s_I(\varepsilon_L) = s_H(\varepsilon_L)$. Moreover, the transition occurs at the energy where the entropies of the low energy branch and of the high energy branch are equal, i.e., that $s_L(\varepsilon_t) = s_H(\varepsilon_t)$. We then obtain that $A_1 = A_2$. It should remarked that the values of the three branches of $\beta(\varepsilon)$ at ε_t determine the size of the temperature jump. Indeed, $\beta_L(\varepsilon_t) = \beta_L^*$, $\beta_I(\varepsilon_t) = \beta_I^*$ and $\beta_H(\varepsilon_t) = \beta_H^*$ (see Fig. 16).

The equal area condition implies that

$$\int_{\beta_L}^{\beta_H} \mathrm{d}\beta\, [\varepsilon(\beta) - \varepsilon_t] = 0. \tag{113}$$

Using $\varepsilon = \mathrm{d}\phi/\mathrm{d}\beta$, one gets

$$\phi(\beta_H) - \phi(\beta_L) - \varepsilon_t(\beta_H - \beta_L) = 0, \tag{114}$$

which, after defining the *generalized entropy*

$$\widehat{s}(\beta, \varepsilon) = \beta\varepsilon - \phi(\beta), \tag{115}$$

leads to

$$\widehat{s}(\beta_L, \varepsilon_t) = \widehat{s}(\beta_H, \varepsilon_t), \tag{116}$$

which is the condition equivalent to (69).

The first time Maxwell's construction appears for self-gravitating systems is in Ref. [112]. It was later extended to microcanonical phase transitions in Refs. [113,114]. On the other hand canonical and microcanonical caloric curves were studied in Ref. [115]. It is also possible that the curve is made of several disconnected branches. We then emphasize that for more complex caloric curves than the one in Fig. 16, the evaluation of Maxwell's areas must be done cautiously [4,34].

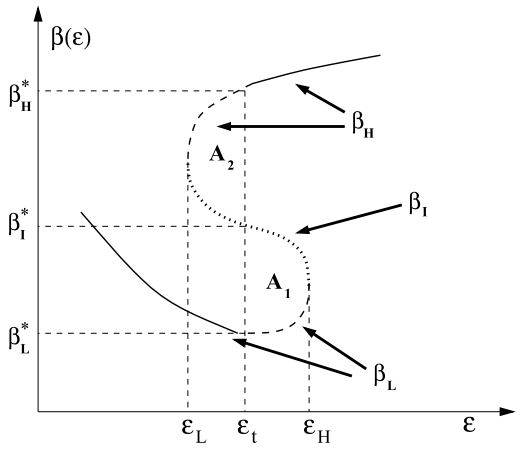


Fig. 16. Typical shape of the $\beta(\varepsilon)$ curve at a microcanonical first order phase transition. The transition energy ϵ_t is determined by an equal area $A_1 = A_2$ Maxwell's construction. All states are represented: stable (solid line), metastable (dashed lines) and unstable (dotted line). The inverse temperature jump is given by $\beta_H^* - \beta_L^*$.

4.3.3. Negative susceptibility

We have seen that ensemble inequivalence can give rise to negative specific heat in the microcanonical ensemble. We show here that another consequence of ensemble inequivalence is the existence of equilibrium microcanonical states with a negative magnetic susceptibility, a non-negative quantity in the canonical ensemble. We will follow a treatment close to that of Section 4.3.1.

Fixing the energy and the magnetization, the entropy is given by Eq. (95). From the first principle of thermodynamics, that for magnetic systems reads $T\mathrm{d}S = \mathrm{d}E - h\mathrm{d}M$, with E and M the internal energy and the total magnetization of the system, respectively, it is straightforward to prove the following formula for the average effective magnetic field in the microcanonical ensemble (see also [116])

$$h(\varepsilon, m) = -\frac{\frac{\partial \tilde{s}}{\partial m}}{\frac{\partial \tilde{s}}{\partial \varepsilon}} = -\frac{1}{\beta(\varepsilon, m)}\frac{\partial \tilde{s}}{\partial m}. \tag{117}$$

Taking into account that, like β is canonically conjugated to H, βh is canonically conjugated to m, it is natural to define the following partition function

$$Z(\beta, h, N) = \sum_{\{S_1,\ldots,S_N\}} \exp\left\{-\beta\left[H(\{S_i\}) - hM(\{S_i\})\right]\right\}, \tag{118}$$

where, as in Eq. (95), $M(\{S_i\})$ is the total magnetization corresponding to configuration $\{S_i\}$. Analogously to Eq. (99), the free energy is

$$f(\beta, h) = -\frac{1}{\beta}\lim_{N\to\infty}\frac{1}{N}\ln Z(\beta, h, N) = \inf_{\varepsilon, m}\left[\varepsilon - hm - \frac{1}{\beta}\tilde{s}(\varepsilon, m)\right]. \tag{119}$$

For $h = 0$, we obviously recover Eq. (99). As in Section 4.3.1, we see that the relation between the two ensembles can be studied by analyzing the single function $\tilde{s}(\varepsilon, m)$. Taking into account Eq. (117), the variational problem in Eq. (119), together with the variational problem that defines $s(\varepsilon)$ can be solved by imposing that

$$\frac{\partial \tilde{s}}{\partial m} = -h\beta \tag{120}$$

$$\frac{\partial \tilde{s}}{\partial \varepsilon} = +\beta, \tag{121}$$

that generalize the conditions (100) and (101) to $h \neq 0$, providing the functions $\varepsilon(\beta, h)$ and $m(\beta, h)$. Independently of the value of h, the stability conditions in the canonical ensemble are the same as those given in Eqs. (104) and (105). In the microcanonical ensemble, since we are not maximizing with respect to m, we have no condition on the second derivative of $\tilde{s}(\varepsilon, m)$ with respect to m.

Magnetic susceptibility is defined as

$$\chi = \frac{\partial m}{\partial h}. \tag{122}$$

Deriving (120) and (121) with respect to h, one obtains

$$\tilde{s}_{m\varepsilon}\frac{\partial\varepsilon}{\partial h}+\tilde{s}_{mm}\frac{\partial m}{\partial h}=-\beta, \tag{123}$$

$$\tilde{s}_{\varepsilon\varepsilon}\frac{\partial\varepsilon}{\partial h}+\tilde{s}_{\varepsilon m}\frac{\partial m}{\partial h}=0, \tag{124}$$

from which we get

$$\chi=-\beta\frac{\tilde{s}_{\varepsilon\varepsilon}}{\tilde{s}_{\varepsilon\varepsilon}\tilde{s}_{mm}-\tilde{s}_{\varepsilon m}^2}. \tag{125}$$

This formula is valid in both the canonical and microcanonical ensemble. However, the results can differ in the two ensembles because the quantities in this formula are computed at different stationary points in the two ensembles. In the canonical ensemble, χ is positive definite for all h because of the stability conditions (104) and (105).

As for the specific heat, the positivity of magnetic susceptibility in the canonical ensemble can be derived on general grounds from Eq. (118), since it is easily shown that susceptibility is proportional to the canonical expectation value $\langle(M-\langle M\rangle)^2\rangle$.

On the other hand, in the microcanonical ensemble, as already remarked, no condition on the second derivatives of $\tilde{s}$ with respect to m is required, and therefore susceptibility can have either sign. Indeed, in Ref. [116] it is shown that a simple ϕ^4 model can display a negative microcanonical susceptibility (see also Section 4.5.5).

With ensemble equivalence, as always happens in short-range systems, magnetic susceptibility is positive also in the microcanonical ensemble. We note that this result is also a byproduct of the convexity property discussed in Section 2.3, since the attainable region in the (ε, m) plane is necessarily convex for short-range systems. In fact, this implies that Eqs. (104) and (105) are satisfied for all equilibrium values (ε, m).

4.4. The Hamiltonian mean field model

Very few examples are known in statistical mechanics where one can explicitly compute microcanonical entropy in cases where the variables are continuous. Everybody knows the perfect gas derivation of microcanonical entropy [10]. However, as soon as one considers interactions, the task becomes unfeasible. On the other hand, since for short-range systems microcanonical and canonical ensemble are equivalent and it is much simpler to perform integrals with Boltzmann weights rather than with Dirac delta functions, much more attention has been devoted to compute free energies. An exception is the study of gravitational systems for which, since it was clarified that ensembles can be non-equivalent [11,12], some attention has been devoted to calculations in different ensembles [3,4,33,34,117–119]. In particular mean-field models have been shown to be solvable using saddle point methods [117,118] and phase transitions in the grand-canonical, canonical and microcanonical ensembles have been studied [120]. On another line of research, simplified gravitational one-dimensional models have been considered [121,122], of which microcanonical solutions have been found [123,124].

This section is devoted to the discussion of a mean-field model, the Hamiltonian Mean Field (HMF) model, whose potential keeps only the first mode of the Fourier expansion of the potential of one-dimensional gravitational and charged sheet models. Indeed, it turns out that the model is nothing but the mean-field version of the XY model (see e.g. Chapter 6.1 of Ref. [125]), which, however, had never been studied in the microcanonical ensemble.

The HMF model has been extensively studied for more than a decade. The simple mean-field interaction allows us to perform analytical calculations, but maintains several complex features of long-range interactions.

It has been shown that the behavior of certain wave-particle Hamiltonians can be understood using the HMF model as a reference. For instance, some equilibrium and non-equilibrium properties of the HMF model can be mapped onto those of the Colson–Bonifacio model of a single-pass Free Electron Laser [126,127].

In this section, we will derive both the canonical and the microcanonical solutions, that we anticipate to be equivalent. The microcanonical solution will be obtained by three different methods: a straightforward one, inspired by the solution of gravitational models (Section 4.4.3), and two more involved ones, using a variational procedure (Section 4.4.4) or large deviations (Section 4.5.3). This latter, although more complicated, will allow us to introduce a method of solution which has a much wider range of applications.

Moreover, this model will serve as a paradigm for discussing important dynamical behaviors typical of long-range interacting systems, which will be presented in Section 5.1.

4.4.1. Introduction

The Hamiltonian Mean Field model [128–130] is defined by the following Hamiltonian

$$H_N=\sum_{i=1}^{N}\frac{p_i^2}{2}+\frac{J}{2N}\sum_{i,j}\left[1-\cos(\theta_i-\theta_j)\right], \tag{126}$$

where $\theta_i \in [0, 2\pi[$ is the position (angle) of the i-th unit mass particle on a circle and p_i the corresponding conjugated momentum. This system can be seen as representing particles moving on a unit circle interacting via an infinite range attractive ($J > 0$) or repulsive ($J < 0$) cosine potential or, alternatively, as classical XY-rotors with infinite range ferromagnetic ($J > 0$) or antiferromagnetic ($J < 0$) couplings. The renormalization factor N of the potential energy is kept not only for historical reasons, but also because, as we have explained in Section 2.1, in this way the energy per particle and temperature are well defined in the $N \to \infty$ limit. In the literature, some authors have treated the case in which the energy is not extensive, i.e. they remove the factor $1/N$. This leads to different thermodynamic limit behaviors [131,132].

Historically, this model has been independently introduced in the continuous time version in Refs. [133–137], and had been previously considered in its time discrete version [138].

We will solve the HMF model in both the canonical and microcanonical ensembles. Alternatively, the model can be solved using the maximum entropy principle for the single particle distribution function [134], which is however suitable only for mean field models.

It is very useful to rewrite the Hamiltonian (126) in a different form, using the definition of the x and y components of the microscopic magnetization

$$m_x = \frac{1}{N}\sum_{i=1}^{N}\cos\theta_i \quad \text{and} \quad m_y = \frac{1}{N}\sum_{i=1}^{N}\sin\theta_i. \tag{127}$$

We then easily find that

$$H_N = \sum_{i=1}^{N}\frac{p_i^2}{2} + \frac{NJ}{2}\left(1 - m^2\right). \tag{128}$$

In the following we will treat only the ferromagnetic case and we will set $J = 1$ without loss of generality. The antiferromagnetic case is less interesting for what equilibrium properties are concerned (the homogenous state is stable at all energies), but it displays interesting dynamical features, like the formation of collective modes under the form of "biclusters" [139–143].

4.4.2. The canonical solution

The canonical solution of this model can be easily derived. After the trivial Gaussian integration over the momenta, the canonical partition function reads

$$Z(\beta, N) = \exp\left(-\frac{N\beta}{2}\right)\left(\frac{2\pi}{\beta}\right)^{N/2}\int \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_N \exp\left\{\frac{\beta}{2N}\left[\left(\sum_{i=1}^{N}\cos\theta_i\right)^2 + \left(\sum_{i=1}^{N}\sin\theta_i\right)^2\right]\right\}. \tag{129}$$

Using the Hubbard–Stratonovich transformation, see (75), this expression becomes

$$Z(\beta, N) = \exp\left(-\frac{N\beta}{2}\right)\left(\frac{2\pi}{\beta}\right)^{N/2}\frac{N\beta}{2\pi}\int \mathrm{d}x_1\mathrm{d}x_2 \exp\left\{N\left[-\frac{\beta(x_1^2 + x_2^2)}{2} + \ln I_0(\beta(x_1^2 + x_2^2)^{\frac{1}{2}})\right]\right\}, \tag{130}$$

where $I_0(z)$ is the modified Bessel function of order 0

$$I_0(z) = \int_0^{2\pi}\mathrm{d}\theta \exp\left(z_1\cos\theta + z_2\sin\theta\right) = \int_0^{2\pi}\mathrm{d}\theta \exp\left(z\cos\theta\right), \tag{131}$$

where $z \equiv \left(z_1^2 + z_2^2\right)^{1/2}$. We can go to polar coordinates in the (x_1, x_2) plane, to obtain

$$Z(\beta, N) = \exp\left(-\frac{N\beta}{2}\right)\left(\frac{2\pi}{\beta}\right)^{N/2}N\beta\int_0^{\infty}\mathrm{d}x \exp\left\{N\left[-\frac{\beta x^2}{2} + \ln I_0(\beta x)\right]\right\}. \tag{132}$$

In the thermodynamic limit $N \to \infty$, the integral in (132) can be computed using the saddle point method, so that the rescaled free energy per particle is

$$\phi(\beta) = \beta f(\beta) = \frac{\beta}{2} - \frac{1}{2}\ln 2\pi + \frac{1}{2}\ln\beta + \inf_{x\geq 0}\left[\frac{\beta x^2}{2} - \ln I_0(\beta x)\right]. \tag{133}$$

The extremal problem in x in the last equation can be easily solved graphically, by looking for the solution of the equation

$$x = \frac{I_1(\beta x)}{I_0(\beta x)}, \tag{134}$$

where $I_1(z)$ is the modified Bessel function of order 1, which is also the derivative of $I_0(z)$. The graphical solution is made easier by the fact that, for real $z > 0$, the function I_1/I_0 is positive, monotonically increasing and with a negative second

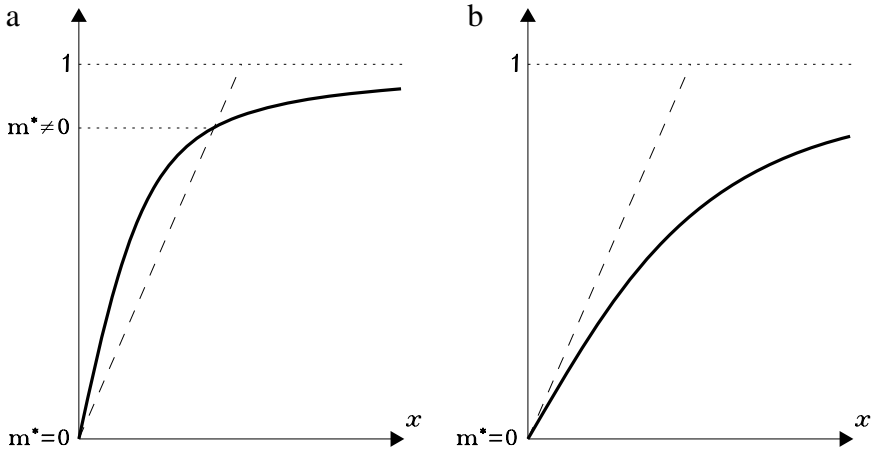


Fig. 17. Graphical solution of Eq. (134) for two values of β, one above and one below the critical value $\beta_c = 2$. In both plots the straight dashed line is the bisectrix, while the curved solid line is the ratio of modified Bessel functions on the r.h.s. of Eq. (134). (a) $\beta = 4$: the equilibrium solution is $m^* \neq 0$; (b) $\beta = 1.5$: the only and equilibrium solution is $m^* = 0$.

derivative (see, e.g., Ref. [144], where a generalization of the HMF model is studied, and where these properties are proved for an entire class of functions that include I_1/I_0). Fig. 17 shows the graphical solution of (134) for two different values of β. For $\beta \leq 2$ the solution is given by $x = m^* = 0$, while for $\beta \geq 2$ the solution monotonically increases with β, approaching $m^* = 1$ for $\beta \to \infty$. The solution $m^* = 0$ of (134), present for all values of β, is not acceptable for $\beta > 2$, since it is not a minimum in formula (133).

It is not difficult to show that the value m^* realizing the extremum in Eq. (133) is equal to the spontaneous magnetization. We emphasize that values of x different from this are not related to the magnetization of a state different from equilibrium state, but that only the minimizer m^*, which is always between 0 and 1, is the magnetization of the equilibrium state. This should be suggested also by the fact that the integration range in Eq. (132) goes from 0 to ∞.

From the above procedure it is clear that the spontaneous magnetization is defined only in its modulus m^*, while there is a continuous degeneracy in its direction.

In conclusion, we have shown that the HMF model displays a second order phase transition at $\beta_c = 2$ ($T_c = 0.5$). The derivative of the rescaled free energy with respect to β gives the energy per particle

$$\varepsilon(\beta) = \frac{1}{2\beta} + \frac{1}{2} - \frac{1}{2}(m^*(\beta))^2. \tag{135}$$

As already evident from the Hamiltonian, the lower bound of ε is 0. At the critical temperature the energy is $\varepsilon_c = \frac{3}{4}$. A plot of the function $\phi(\beta)$ is shown in the next subsection, after the microcanonical solution.

4.4.3. The microcanonical solution

Since we have shown that the HMF model has a second order phase transition in the canonical ensemble, we could immediately conclude, according to the remark in Section 4.1.3, that ensembles are equivalent for this model. Deriving the entropy would be straightforward using the inverse Legendre–Fenchel transform of the free energy computed in the previous subsection (see formula (133)). However, since in the following we will need the explicit expression of the microcanonical entropy of the HMF model both as a function of the energy and of magnetization in order to solve a generalized HMF model in the microcanonical ensemble, we will devote this subsection, the following and Section 4.5.3 to the derivation of the entropy for the HMF model using several methods. The reason for such a thorough derivation is mainly pedagogical: we want to show the application of three different methods to a simple model. The first method is borrowed from similar ones used in the context of gravitational systems. The second one, more general, is a variational method based on a strong hypothesis on the form the partition function, and allows a straightforward derivation of the entropy, also in cases of ensemble inequivalence. The third method illustrates the application of large deviation theory, which is the most general currently available tool to solve systems with long-range interactions.

The microcanonical solution has been heuristically obtained, under the hypothesis of concave entropy, in Ref. [145] and in a different form in Ref. [146].

The simplicity of the Hamiltonian makes it possible to obtain directly the thermodynamic limit of the entropy per particle, as we now show. First of all, let us introduce the method usually applied in self-gravitating systems [22]. Denoting by K and U the kinetic and the potential energy, respectively, the number of microscopic configurations corresponding to the energy

E in a generic system is given by

$$\Omega(E,N) = \int \prod_i \mathrm{d}p_i \mathrm{d}\theta_i \delta(E - H_N) \tag{136}$$

$$= \int \prod_i \mathrm{d}p_i \mathrm{d}\theta_i \underbrace{\int \mathrm{d}K\, \delta\left(K - \sum_i \frac{p_i^2}{2}\right)}_{=1} \delta\left(E - K - U(\{\theta_i\})\right) \tag{137}$$

$$= \int \mathrm{d}K \underbrace{\int \prod_i \mathrm{d}p_i \delta\left(K - \sum_i \frac{p_i^2}{2}\right)}_{\Omega_{\mathrm{kin}}(K)} \underbrace{\int \prod_i \mathrm{d}\theta_i \delta\left(E - K - U(\{\theta_i\})\right)}_{\Omega_{\mathrm{conf}}(E-K)} . \tag{138}$$

The factor Ω_{kin}, which is related to the surface of the hypersphere with radius $R = \sqrt{2K}$ in N dimensions, can be computed straightforwardly using the properties of the Dirac δ function, obtaining the expression: $\Omega_{\mathrm{kin}} = 2\pi^{N/2} R^{N-2}/\Gamma(N/2)$. Using the asymptotic expression of the Γ-function, $\ln\Gamma(N) \simeq (N - 1/2)\ln N - N + (1/2)\ln(2\pi)$, and keeping only the terms that do not give a vanishing contribution to the entropy per particle in the thermodynamic limit, one obtains

$$\Omega_{\mathrm{kin}}(K) \overset{N\to+\infty}{\sim} \exp\left(\frac{N}{2}\left[1 + \ln\pi - \ln\frac{N}{2} + \ln(2K)\right]\right) \tag{139}$$

$$= \exp\left(\frac{N}{2}\left[1 + \ln(2\pi) + \ln u\right]\right), \tag{140}$$

where $u = 2K/N$. Defining the configurational entropy per particle $s_{\mathrm{conf}}(\tilde{u}) = (\ln\Omega_{\mathrm{conf}}(N\tilde{u}))/N$, where $\tilde{u} = U/N = (E - K)/N = \varepsilon - u/2$, Eq. (138) can be rewritten as

$$\Omega(N\varepsilon, N) \overset{N\to+\infty}{\sim} \frac{N}{2}\int \mathrm{d}u \exp\left[N\left(\frac{1}{2} + \frac{\ln(2\pi)}{2} + \frac{1}{2}\ln u + s_{\mathrm{conf}}(\tilde{u})\right)\right]. \tag{141}$$

Hence, solving the integral in the saddle point approximation, we obtain the entropy

$$s(\varepsilon) = \lim_{N\to+\infty} \frac{1}{N}\ln\Omega_N(\varepsilon N) \tag{142}$$

$$= \frac{1}{2} + \frac{1}{2}\ln(2\pi) + \sup_u\left[\frac{1}{2}\ln u + s_{\mathrm{conf}}(\tilde{u})\right]. \tag{143}$$

We note that this expression is quite general, in the sense that it is valid for any system in which the kinetic energy assumes the usual quadratic form. To proceed further, we need an explicit expression for the configurational entropy s_{conf}, something which is generally not easily feasible.

Now we use the fact that the potential energy of the HMF model, as evident in (128), is a very simple function of the microscopic magnetization $\mathbf{m} = (m_x, m_y)$, with a one to one correspondence between the value of the potential energy U and the modulus of the microscopic magnetization $m^2 = m_x^2 + m_y^2$. In fact, if we define

$$\Omega_m(m) = \int \prod_i \mathrm{d}\theta_i \delta\left(\sum_i \cos\theta_i - Nm\right)\delta\left(\sum_i \sin\theta_i\right), \tag{144}$$

we have that this function will be proportional to $\Omega_{\mathrm{conf}}(U)$ for $\tilde{u} = U/N = (1/2 - m^2/2)$. The coefficient of proportionality will give a vanishing contribution to $s_{\mathrm{conf}}(\tilde{u})$ in the thermodynamic limit. We note that, as in the canonical case, there is a continuous degeneracy on the direction of the spontaneous magnetization; therefore, we do not lose generality by choosing the spontaneous magnetization in the direction of the x axis. The integral in (144) can be computed using the Fourier representation of the δ function. We therefore have

$$\Omega_m(m) = \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty}\mathrm{d}q_1 \int_{-\infty}^{\infty}\mathrm{d}q_2 \int \prod_i \mathrm{d}\theta_i \exp\left[\mathrm{i}q_1\left(\sum_i \cos\theta_i - Nm\right)\right]\exp\left[\mathrm{i}q_2\left(\sum_i \sin\theta_i\right)\right] \tag{145}$$

$$= \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty}\mathrm{d}q_1 \int_{-\infty}^{\infty}\mathrm{d}q_2 \exp\left\{N\left[-\mathrm{i}q_1 m + \ln J_0((q_1^2 + q_2^2)^{\frac{1}{2}})\right]\right\}, \tag{146}$$

where $J_0(z)$ is the Bessel function of order 0

$$J_0(z) = \int_0^{2\pi}\mathrm{d}\theta \exp(\mathrm{i}z\cos\theta). \tag{147}$$

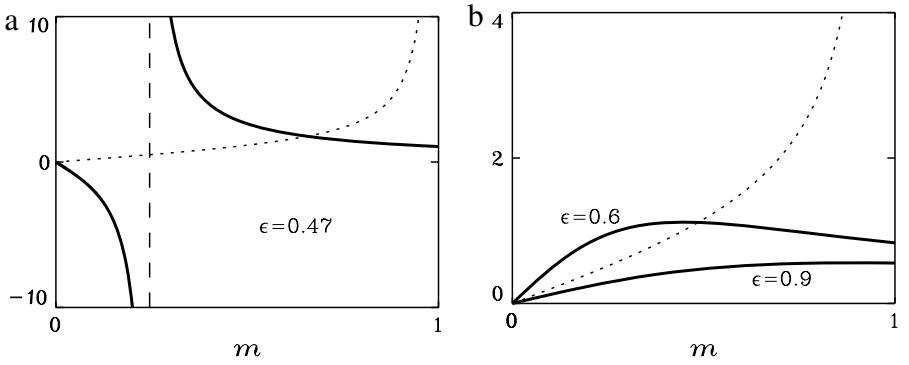


Fig. 18. Graphical solution of Eq. (152) for three values of the energy ε, one above and two below the critical value $\varepsilon_c = 3/4$. In both plots the dotted curve is the function $B_{\rm inv}(m)$, while the solid curve is the function of the first term in Eq. (152). This last function diverges for $m^2 = 1 - 2\varepsilon$, which is between 0 and 1 for $0 \leq \varepsilon \leq 1/2$. The corresponding asymptote is the vertical dotted line in the left panel. (a) $\varepsilon = 0.47$: the relevant solution is the one with $m > 0$; (b) $\varepsilon = 0.6$ and $\varepsilon = 0.9$: in the first case the relevant solution is the one with $m > 0$, while in the second case the only solution is $m = 0$.

To solve the integral in (145) with the saddle point method, we have to consider q_1 and q_2 as complex variables. Using that the derivative of J_0 is $-J_1$, the opposite of the Bessel function of order 1, the saddle point has to satisfy the following equations

$$\begin{aligned} &-\mathrm{i}m - \frac{J_1}{J_0}((q_1^2+q_2^2)^{\frac{1}{2}})\frac{q_1}{(q_1^2+q_2^2)^{\frac{1}{2}}} = 0 \\ &-\frac{J_1}{J_0}((q_1^2+q_2^2)^{\frac{1}{2}})\frac{q_2}{(q_1^2+q_2^2)^{\frac{1}{2}}} = 0. \end{aligned} \tag{148}$$

The solution of these equations is $q_2 = 0$ and $q_1 = -\mathrm{i}\gamma$, where γ is the solution of the equation

$$\frac{I_1(\gamma)}{I_0(\gamma)} = m. \tag{149}$$

Here, we have used the properties $J_0(\mathrm{i}z) = I_0(z)$ and $J_1(\mathrm{i}z) = \mathrm{i}I_1(z)$. Denoting by $B_{\rm inv}$ the inverse function of I_1/I_0, we can write in the thermodynamic limit

$$s_{\rm conf}\left(\frac{1}{2} - \frac{1}{2}m^2\right) = \lim_{N\to+\infty}\frac{1}{N}\ln\Omega_m(m) = -mB_{\rm inv}(m) + \ln I_0(B_{\rm inv}(m)). \tag{150}$$

We can now substitute (150) in (143), using that $u = 2(\varepsilon - 1/2 + m^2/2)$, and performing equivalently a maximization over m instead of that over u, to obtain

$$s(\varepsilon) = \frac{1}{2} + \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln 2 + \sup_{m\geq m_0}\left[\frac{1}{2}\ln\left(\varepsilon - \frac{1}{2} + \frac{1}{2}m^2\right) - mB_{\rm inv}(m) + \ln I_0(B_{\rm inv}(m))\right], \tag{151}$$

with $m_0^2 = \sup[0, 1-2\varepsilon]$. The maximization problem over m is solved graphically, looking for the solutions of the equation

$$\frac{m}{2\varepsilon - 1 + m^2} - B_{\rm inv}(m) = 0. \tag{152}$$

The graphical solution $m = m(\varepsilon)$ is shown in Fig. 18, and it gives the following results. For $0 \leq \varepsilon \leq 3/4$, the magnetization $m(\varepsilon)$ monotonically decreases from 1 to 0, while for $\varepsilon > 3/4$ the solution is always $m = 0$. At $\varepsilon = 3/4$, there is a second order phase transition, a first signature that the two ensembles give equivalent predictions.

Indeed, it is easy to prove that $s(\varepsilon)$ is concave. First, we easily derive from (151) that the inverse temperature $\beta(\varepsilon)$ is

$$\beta(\varepsilon) = \frac{\mathrm{d}s}{\mathrm{d}\varepsilon} = \frac{1}{2\varepsilon - 1 + m^2(\varepsilon)}. \tag{153}$$

If $\beta(\varepsilon)$ has a negative derivative, then $s(\varepsilon)$ is concave. The negativity of this derivative is trivial for $\varepsilon > 3/4$, when $m(\varepsilon) = 0$. For $\varepsilon \leq 3/4$ we can proceed as follows. We note that, using the last equation, we can write (152) also as

$$m = \frac{I_1(m\beta(\varepsilon))}{I_0(m\beta(\varepsilon))}. \tag{154}$$

We have just proved graphically that (152), and thus (154), have a unique solution for each ε, that decreases if ε increases. Studying (134) in the canonical ensemble, we have found that this unique solution has the property that m decreases when

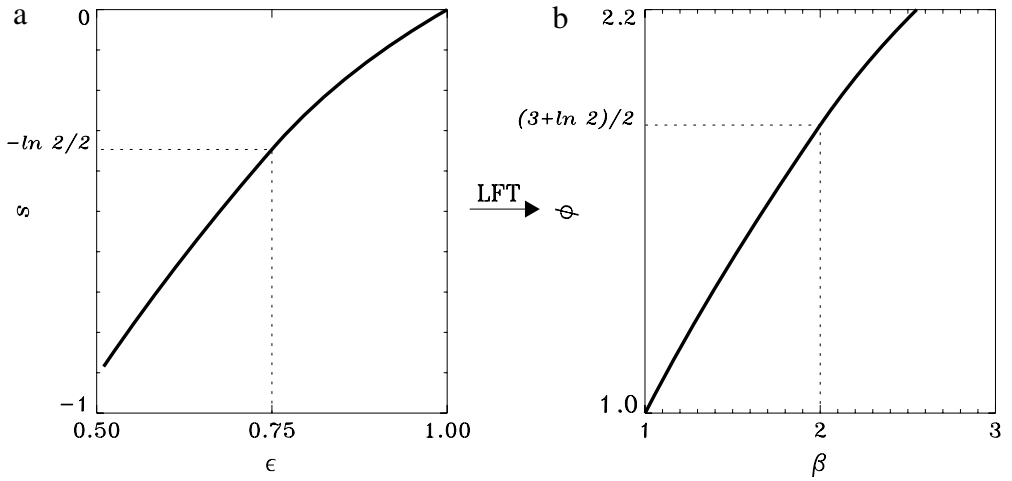


Fig. 19. Entropy versus energy (a) and rescaled free energy versus inverse temperature (b) for the HMF model (126) with $J = 1$. The dotted lines are traced at the phase transition point.

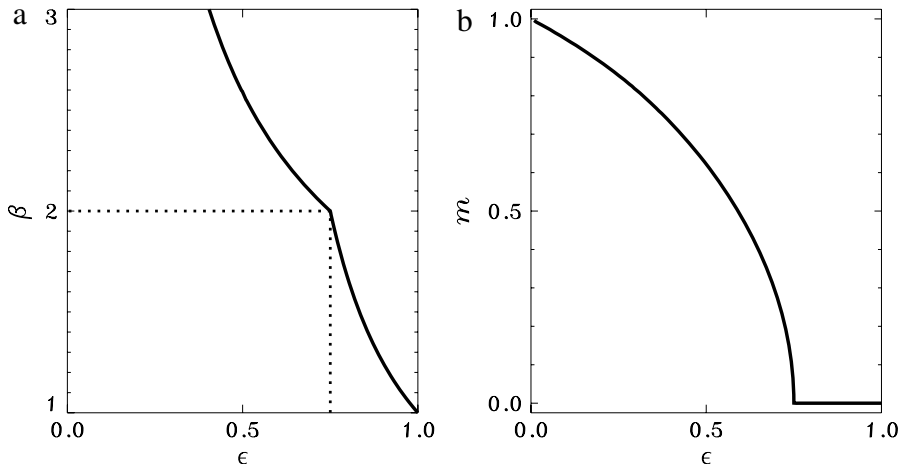


Fig. 20. Inverse temperature versus energy (a) and magnetization versus energy (b) for the HMF model (126) with $J = 1$. The dotted lines are traced at the phase transition point.

β decreases. Therefore, in the present case an increase in ε results in a decrease in $\beta(\varepsilon)$, also when $m(\varepsilon) > 0$. This proves the concavity of the entropy $s(\varepsilon)$ and therefore ensemble equivalence. The concavity property of the microcanonical entropy (151) ensures that it could also be computed by the Legendre–Fenchel transform of ϕ, the rescaled free energy computed in the canonical ensemble and reported in Eq. (133).

We conclude this section devoted to the explicit calculation of the canonical and microcanonical solution of the HMF model by showing the plots of the relevant thermodynamic variables. In Fig. 19 we show the full dependence of the entropy on the energy (see Eq. (151)) and the rescaled free energy versus the inverse temperature (see Eq. (133)). In Fig. 20 we plot the caloric curve (see Eq. (153)) and the dependence of the order parameter on the energy, obtainable from Eq. (152).

4.4.4. Min–max procedure

We have already emphasized that the microcanonical partition function is much more difficult to obtain than the canonical one. Although we have been able to obtain it by a direct counting for the BEG model, this procedure is applicable only for discrete variables. In this subsection we will show how the microcanonical entropy can be obtained when the canonical free energy has been derived through an optimization procedure of the type shown in (133). At the same time, this discussion will allow us to understand how ensemble inequivalence can arise. However, it should be remarked that the derivation presented here is not rigorous and relies on strong assumptions. We will follow the argument presented in Ref. [147], and later generalized in Ref. [148], although we will give here a somewhat different proof of the result.

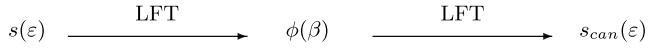


Fig. 21. Relation between microcanonical entropy $s(\varepsilon)$, rescaled free energy $\phi(\beta)$ and canonical entropy $s_{\rm can}(\varepsilon)$. LFT indicates the Legendre–Fenchel transform.

It is crucial to assume that the canonical partition function, similarly to what has been done in Eq. (132) for the HMF model, is given by the following integral

$$Z(\beta, N) = \int_{-\infty}^{+\infty} \mathrm{d}x \exp\left[-N\tilde{\phi}(\beta, x)\right], \tag{155}$$

with $\tilde{\phi}(\beta, x)$ a differentiable function of β for ($\beta \geq 0$) and a dummy variable x. This is a crucial assumption, which is at the core of the applicability of the method. We do not explicitly give any clue to what is the nature and physical meaning of the variable x, but of course we have in mind the integration variable which appears in the Hubbard–Stratonovich transformation. Moreover, it should be observed that the variable x has nothing to do with the order parameter m. Even the variation range of these two variables is different: in principle the range of x is the full real axis and the integral in Eq. (155) is required to converge (for example, in the Hubbard–Stratonovich trasformation there is a dependence of the integrand of the kind $\exp(-x^2)$).

In the thermodynamic limit, we therefore have

$$\phi(\beta) = \inf_x \tilde{\phi}(\beta, x). \tag{156}$$

The *canonical entropy* is defined by the following variational principle,

$$s_{\rm can}(\varepsilon) = \inf_{\beta \geq 0} [\widehat{s}(\beta, \varepsilon)] \tag{157}$$

where $\widehat{s}$ is defined in formula (115). This is nothing but the Legendre–Fenchel transform of $\phi(\beta)$. We observe that this expression is valid also in the case in which ε is such that there is no β value for which the derivative of $\phi(\beta)$ is equal to ε. In our case, we can insert Eq. (156) to obtain

$$s_{\rm can}(\varepsilon) = \inf_{\beta \geq 0} \left\{ \sup_x \left[\beta\varepsilon - \tilde{\phi}(\beta, x)\right] \right\}. \tag{158}$$

One should remark that this relation is valid only if the assumption made for the partition function in Eq. (155) is fulfilled. In particular the x-dependence of $\tilde{\phi}$ cannot be neglected in order for the integral in Eq. (155) to converge. The relation between $s(\varepsilon)$, $\phi(\beta)$ and $s_{\rm can}(\varepsilon)$ is summarized in Fig. 21.

When ensembles are equivalent $s_{\rm can}(\varepsilon) = s(\varepsilon)$. We will show here that a necessary and sufficient condition for the ensembles to be inequivalent is that $s(\varepsilon)$ is strictly smaller than $s_{\rm can}(\varepsilon)$. This is for instance what happens for the BEG model, for which $s_{\rm can}(\varepsilon)$ is the concave envelope of $s(\varepsilon)$, as shown in Fig. 13.

We will prove here that, once the function $\tilde{\phi}(\beta, x)$ is known, one can introduce its Legendre–Fenchel transform

$$\tilde{s}(\varepsilon, x) = \inf_\beta \left[\beta\varepsilon - \tilde{\phi}(\beta, x)\right] \tag{159}$$

and obtain microcanonical entropy by the following formula

$$s(\varepsilon) = \sup_x [\tilde{s}(\varepsilon, x)] = \sup_x \left\{ \inf_{\beta \geq 0} \left[\beta\varepsilon - \tilde{\phi}(\beta, x)\right] \right\}. \tag{160}$$

The only difference in expressions (158) and (160) is just the order in which the minimum with respect to β and the maximum with respect to x are taken. Although this might seem a detail, it determines a different result (see Appendix A, where this is proven in full generality). It turns out that

$$s(\varepsilon) \leq s_{\rm can}(\varepsilon). \tag{161}$$

Let us remark that the entropy defined in formula (159) does not coincide with $\tilde{s}(\varepsilon, m)$. In particular, one can easily verify that $\tilde{s}(\varepsilon, x)$ is always concave in ε for all x values, since it is obtained from a Legendre–Fenchel transform. On the contrary, $\tilde{s}(\varepsilon, m)$ can be non-concave in ε.

Let us briefly sketch the proof of formula (160), which is based on the analysis of microcanonical partition function. Using the Laplace representation of the Dirac delta function, this latter can be expressed as

$$\Omega(E, N) = \sum_{\{S_1, \ldots, S_N\}} \delta(E - H(\{S_i\})) \tag{162}$$

$$= \frac{1}{2\pi \mathrm{i}} \int_{\beta - \mathrm{i}\infty}^{\beta + \mathrm{i}\infty} \mathrm{d}\lambda \sum_{\{S_1, \ldots, S_N\}} \exp[\lambda(E - H(\{S_i\}))] \tag{163}$$

$$= \frac{1}{2\pi \mathrm{i}} \int_{\beta - \mathrm{i}\infty}^{\beta + \mathrm{i}\infty} \mathrm{d}\lambda\, Z(\lambda, N) \exp(\lambda E) \tag{164}$$

where $\beta = \mathrm{Re}(\lambda) > 0$ is the inverse temperature. We use λ, instead of β, as an integration variable, because we are considering the analytical continuation of $Z(\lambda, N)$ to the complex plane. The last integral cannot be solved, in the thermodynamic limit, using the saddle point method because, after expressing $Z(\beta, N) \sim \exp[-N\beta f(\beta)]$, the function $\beta f(\beta)$ is not in general differentiable for all β. In spite of this, we can heuristically argue that the integral will be dominated by the value of the integrand at a real value of λ ($\lambda = \beta + \mathrm{i}\lambda_I$), otherwise we would obtain an oscillatory behavior of $\Omega(E, N)$, giving rise to negative values, which are impossible for the density of states. To have a proof of this we can proceed as follows.

We are assuming to be in cases where $Z(\beta, N)$, for real β, can be expressed as in (155), with $\tilde{\phi}$ analytic. Then, this integral representation will be valid, in the complex λ plane, for at least a strip that includes the real axis, let's say for $|\lambda_I| < \Delta$, with $\Delta > 0$.

We now divide the integral in (164) in three intervals, defined by $\lambda_I < -\delta$, $-\delta < \lambda_I < \delta$ and $\lambda_I > \delta$, respectively, with $0 < \delta < \Delta$. In Appendix B, we show that the contribution to the integral in λ coming from values of λ_I outside the strip, i.e. for values of λ_I with $|\lambda_I| > \Delta$, is exponentially small in N. Therefore the calculation of the microcanonical partition function reduces to performing the following integral

$$\Omega(E, N) = \frac{1}{2\pi\mathrm{i}} \int_{\beta-\mathrm{i}\delta}^{\beta+\mathrm{i}\delta} \mathrm{d}\lambda \mathrm{e}^{N\lambda\varepsilon} Z(\lambda, N) = \frac{1}{2\pi\mathrm{i}} \int_{-\infty}^{+\infty} \mathrm{d}x \int_{\beta-\mathrm{i}\delta}^{\beta+\mathrm{i}\delta} \mathrm{d}\lambda \exp\left(N[\lambda\varepsilon - \tilde{\phi}(\lambda, x)]\right), \tag{165}$$

where $0 < \delta < \Delta$, and where, in the second equality, we have used the fact that inside the strip $|\lambda_I| < \Delta$ we can represent $Z(\beta, N)$ as in (155). Since $\tilde{\phi}$ is analytic in all the domain of integration, we can perform the integral in λ in the large N limit using the saddle point method. It can be shown that for each value of x the real part of the argument of the exponential in (165) is larger if computed on the real axis $\lambda_I = 0$. Indeed, we have

$$\int_{-\infty}^{+\infty} \mathrm{d}x \exp\{N[\lambda\varepsilon - \tilde{\phi}(\lambda, x)]\} = \exp(N\lambda\varepsilon) Z(\lambda, N) \tag{166}$$

$$= \exp(N\lambda\varepsilon) \sum_{\{S_1,\ldots,S_N\}} \exp\{-\beta H(\{S_i\}) - \mathrm{i}\lambda_I H(\{S_i\})\} \tag{167}$$

$$= \exp(N\beta\varepsilon) \sum_{\{S_1,\ldots,S_N\}} \exp\{-\beta H(\{S_i\}) - \mathrm{i}\lambda_I H(\{S_i\}) + \mathrm{i}\lambda_I N\varepsilon\} \tag{168}$$

$$= \exp(N\beta\varepsilon) \langle \exp\left(\mathrm{i}\lambda_I [N\varepsilon - H(\{S_i\})]\right)\rangle Z(\beta, N), \tag{169}$$

where in the last expression the average $\langle\cdot\rangle$ is performed with Boltzmann weight $\exp[\beta H(\{S_i\})]$. Using now the definition $Z(\beta, N) = \int \mathrm{d}x \exp[-N\tilde{\phi}(\beta, x)]$, one can rewrite $\Omega(E, N)$ as

$$\Omega(E, N) = \frac{1}{2\pi\mathrm{i}} \int_{-\infty}^{+\infty} \mathrm{d}x \int_{\beta-\mathrm{i}\delta}^{\beta+\mathrm{i}\delta} \mathrm{d}\lambda \exp\left(-N\tilde{\phi}(\beta, x) + N\beta\varepsilon + \ln\langle \exp\left(\mathrm{i}\lambda_I [N\varepsilon - H(\{S_i\})]\right)\rangle\right). \tag{170}$$

The real part of the exponent in the integral is obtained by replacing $\ln(\langle\cdot\rangle)$ with $\ln(|\langle\cdot\rangle|)$. The maximum of this logarithm is obtained for $\lambda_I = 0$, since for all other nonzero values of λ_I, when performing the average in $\langle\cdot\rangle$, one would sum unit vectors $\exp(\mathrm{i}\lambda_I\alpha)$ with different values of the phase α, depending on the specific configuration $\{S_i\}$, obtaining a result which will have certainly a smaller modulus than summing all the vectors in phase with $\lambda_I = 0$. Therefore, if there is a saddle point on the real axis, it will certainly give a larger real part for the argument of the exponential than other saddle points eventually present outside the real axis. Moreover, we have just proven that the saddle is necessarily a maximum along the imaginary λ direction. Therefore, from the general properties of holomorphic functions, it will be a minimum along the $\beta = \mathrm{Re}(\lambda)$ axis. Once the integral over λ is performed, the remaining integral over x can be computed using again the saddle point method, now for a real function, which gives a maximum over x. Combining all this, one gets the formula for the microcanonical entropy (160).

Another way of arguing [147] is to remark that, since $\tilde{\phi}(\lambda, x)$ is obtained by analytically continuing a real function, if saddle points are present in the complex λ plane, they necessarily appear in complex conjugate pairs. This would induce oscillations in the values of $\Omega(E, N)$ as a function of N, which would imply that $\Omega(E, N)$ could even take negative values. This is absurd and leads to exclude the presence of saddle points out of the real λ axis.

We have already remarked that ensemble inequivalence is a consequence of inequality (161), which in turn derives from the different order of the minimum in β with respect to the maximum in x in expressions (158) and (160). Let us now study in more detail the extrema defined by the two different variational problems. The first order stationarity conditions are

$$\frac{\partial\tilde{\phi}}{\partial\beta} = \varepsilon \tag{171}$$

$$\frac{\partial\tilde{\phi}}{\partial x} = 0, \tag{172}$$

which are of course the same for the two ensembles. However, the stability conditions deriving from the two problems are different. For what concerns the canonical entropy $s_{\rm can}(\varepsilon)$, we have the conditions

$$\frac{\partial^2\tilde{\phi}}{\partial x^2} > 0 \tag{173}$$

$$\frac{\partial^2\tilde{\phi}}{\partial\beta^2}\frac{\partial^2\tilde{\phi}}{\partial x^2} - \left(\frac{\partial^2\tilde{\phi}}{\partial\beta\partial x}\right)^2 < 0, \tag{174}$$

while for the microcanonical entropy $s(\varepsilon)$ we have

$$\frac{\partial^2\tilde{\phi}}{\partial\beta^2} < 0 \tag{175}$$

$$\frac{\partial^2\tilde{\phi}}{\partial\beta^2}\frac{\partial^2\tilde{\phi}}{\partial x^2} - \left(\frac{\partial^2\tilde{\phi}}{\partial\beta\partial x}\right)^2 < 0. \tag{176}$$

A necessary and sufficient condition to satisfy (173)–(176) is that $\partial^2\tilde{\phi}/\partial\beta^2 < 0$ and $\partial^2\tilde{\phi}/\partial x^2 > 0$. However, since the conditions are different in the two ensembles, one can find values of β and x that correspond to stable states in one ensemble but are unstable in the other. It should be noted that one could find more than one stationary point in a given ensemble, satisfying the corresponding stability condition. Obviously, in this case one has to choose the global extremum. If the global stable extrema are different in the two ensembles, then we have ensemble inequivalence. Tightly linked to stability is the sign of specific heat. Indeed, using the expression

$$c_V = -\beta^2\frac{\partial\tilde{\phi}(\beta, x(\beta))}{\partial\beta^2}, \tag{177}$$

where $x(\beta)$ is obtained by solving (171) and (172), one can obtain an expression for the specific heat which is valid in both the ensembles

$$c_V = -\beta^2\frac{\frac{\partial^2\tilde{\phi}}{\partial\beta^2}\frac{\partial^2\tilde{\phi}}{\partial x^2} - \left(\frac{\partial^2\tilde{\phi}}{\partial\beta\partial x}\right)^2}{\frac{\partial^2\tilde{\phi}}{\partial x^2}}. \tag{178}$$

We see from (173) and (174) that this expression is positive in the canonical ensemble. However, in the microcanonical ensemble the conditions (175) and (176) do not determine the sign of $\partial^2\tilde{\phi}/\partial x^2$, and thus the specific heat can have either sign in the microcanonical ensemble.

Let us check that, using the method described in this subsection, we can rederive the microcanonical entropy of the HMF model, Eq. (151). For the HMF model, the function $\tilde{\phi}(\beta, x)$ can be obtained from (132)

$$\tilde{\phi}(\beta, x) = \frac{\beta}{2} - \frac{1}{2}\ln 2\pi + \frac{1}{2}\ln\beta + \frac{\beta x^2}{2} - \ln I_0(\beta x). \tag{179}$$

The stationarity conditions (171) and (172) read in this case

$$\frac{\partial\tilde{\phi}}{\partial\beta} = \frac{1}{2} + \frac{1}{2\beta} + \frac{1}{2}x^2 - x\frac{I_1(\beta x)}{I_0(\beta x)} = \varepsilon \tag{180}$$

$$\frac{\partial\tilde{\phi}}{\partial x} = \beta x - \beta\frac{I_1(\beta x)}{I_0(\beta x)} = 0. \tag{181}$$

Inserting the second equation in the first, we find that $\beta^{-1} = 2\varepsilon - 1 + x^2$. Substituting back in the second equation, and using the function $B_{\rm inv}$, the inverse function of I_1/I_0 previously defined, we have

$$B_{\rm inv}(x) - \frac{x}{2\varepsilon - 1 + x^2} = 0, \tag{182}$$

which is identical to Eq. (152). The computation of the second derivatives gives

$$\frac{\partial^2\tilde{\phi}}{\partial\beta^2} = -\frac{1}{2\beta^2} - x\frac{\partial}{\partial x}\left(\frac{I_1(\beta x)}{I_0(\beta x)}\right) < 0 \tag{183}$$

$$\frac{\partial^2\tilde{\phi}}{\partial x^2} = \beta - \beta\frac{\partial}{\partial x}\left(\frac{I_1(\beta x)}{I_0(\beta x)}\right) > 0. \tag{184}$$

These inequalities are both satisfied at the stationary points determined by (180) and (181). Actually, Eq. (183) is identically satisfied, since the derivative of I_1/I_0 is positive definite. This confirms ensemble equivalence for the HMF model.

Finally, Eqs. (158) and (160) give

$$s(\varepsilon) = s_{\rm can}(\varepsilon) = \frac{1}{2} + \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln 2 + \frac{1}{2}\ln\left(\varepsilon - \frac{1}{2} + \frac{1}{2}x^2\right) - \frac{x^2}{2\varepsilon - 1 + x^2} + \ln I_0\left(\frac{x^2}{2\varepsilon - 1 + x^2}\right), \tag{185}$$

with x satisfying Eq. (182). We thus obtain an expression identical to (151), taking into account Eq. (152).

4.5. The large deviation method and its applications

The mathematical theory of large deviations is a field in its own, and obviously it is outside the scope of this review to give details of this theory and to treat at the level of mathematical rigor what will be presented about it. The methods based on large deviation theory and their applications to problems in statistical mechanics have been popularized among theoretical physicists in several books and review papers [99,149,150]. We would like to cite here a few interesting works in which physical systems with long-range interactions have been studied using this method. Michel and Robert [57] successfully used large deviations techniques to rigorously prove the applicability of statistical mechanics to two-dimensional fluid mechanics, proposed earlier [7,56,151]. Ellis et al. [152] pursued the approach of Robert and Sommeria [56] to solve two-dimensional geophysical systems.

Many particle systems with long-range interactions often offer a relatively simple field of application of the theory of large deviations. This is very interesting, especially if one considers that the outcome of the calculation is the entropy function.

The structure of this section is the following. In Section 4.5.1, where we introduce the method, we will show how to obtain the entropy for a class of systems in which the Hamiltonian can be expressed in a way that is often realized in long-range systems. In Section 4.5.2 we will show the simple application to the 3-states Potts model, where the result can be compared with the direct computation of the entropy. This simple discrete system already shows ensemble inequivalence. In Section 4.5.3, we will treat again the HMF model to introduce the application of the method to systems with continuous variables. Interesting applications will be treated in Sections 4.5.4 and 4.5.5, both devoted to systems presenting ensemble inequivalence: a generalized HMF model and the so called ϕ^4 model. Finally, in Section 4.5.6, we will present the equilibrium solution of the Colson–Bonifacio model of a linear Free Electron Laser.

4.5.1. The computation of the entropy for long-range systems

Denoting again collectively with $x \equiv (\{p_i\}, \{q_i\})$ the phase space variables of a Hamiltonian system, let us suppose that the energy per particle $\varepsilon(x) = H(x)/N$ can be expressed as a function of (few) global "mean-fields" $\mu_1(x), \ldots, \mu_n(x)$; i.e., we suppose that it is possible to write

$$\varepsilon(x) = \bar{\varepsilon}(\mu_1(x), \ldots, \mu_n(x)). \tag{186}$$

This is a situation often realized in long-range systems. Actually, one could require that formula (186) be valid only asymptotically for $N \to \infty$, while for large but finite N the right-hand side could contain a remainder $R(x)$ that can be neglected in the thermodynamic limit. However, in mean-field systems representation (186) is exact for all N. With the specification of the microscopic configuration x, we define what is generally indicated as a microstate of the system. On the contrary, specifying that the system is in a state in which the global variables have given values $\mu_1, \ldots, \mu_n$, we are defining what is called a macrostate of the system. Once the macrostate is chosen, the microscopic configuration is not determined, since all x that satisfy $\mu_k(x) = \mu_k$ for $k = 1, \ldots, n$ belong to the same macrostate.

The identification, in a concrete system, of the global variables is the first step in the application of the large deviation method.

The second step will be the computation of the entropy of the different macrostates, i.e., the calculation of the function

$$\bar{s}(\mu_1, \ldots, \mu_n) = \lim_{N\to\infty} \frac{1}{N} \ln \int {\rm d}x \delta(\mu_1(x) - \mu_1) \ldots \delta(\mu_n(x) - \mu_n). \tag{187}$$

Leaving aside for the moment the problem of computing $\bar{s}(\mu_1, \ldots, \mu_n)$, that at first sight does not seem to be any simpler than computing the entropy function $s(\varepsilon)$, it is easy to see how this last function can be obtainable from $\bar{s}(\mu_1, \ldots, \mu_n)$. In fact, we have

$$\int {\rm d}x \delta\left[N(\varepsilon(x) - \varepsilon)\right] = \int {\rm d}x \delta\left[N(\bar{\varepsilon}(\mu_1(x), \ldots, \mu_n(x)) - \varepsilon)\right] \tag{188}$$

$$= \int {\rm d}x {\rm d}\mu_1 \ldots {\rm d}\mu_n \delta(\mu_1(x) - \mu_1) \ldots \delta(\mu_n(x) - \mu_n) \delta\left[N(\bar{\varepsilon}(\mu_1, \ldots, \mu_n) - \varepsilon)\right] \tag{189}$$

$$\overset{N\to+\infty}{\sim} \int {\rm d}\mu_1 \ldots {\rm d}\mu_n \exp\left[N\bar{s}(\mu_1, \ldots, \mu_n)\right] \delta\left[N(\bar{\varepsilon}(\mu_1, \ldots, \mu_n) - \varepsilon)\right]. \tag{190}$$

We therefore have

$$s(\varepsilon) = \sup_{[\mu_1,\ldots,\mu_n|\bar{\varepsilon}(\mu_1,\ldots,\mu_n)=\varepsilon]} \bar{s}(\mu_1,\ldots,\mu_n). \tag{191}$$

The solution of this extremal problem constitutes the third and final step of the computation. We are left with the problem of computing $\bar{s}(\mu_1,\ldots,\mu_n)$, i.e., with the actual implementation of the second step. Let us introduce the following canonical partition function

$$\bar{Z}(\lambda_1,\ldots,\lambda_n) = \int \mathrm{d}x \exp\left[-N\left(\lambda_1\mu_1(x) + \cdots + \lambda_n\mu_n(x)\right)\right]. \tag{192}$$

Few steps completely analogous to those relating $\phi(\beta) = \beta f(\beta)$ to $s(\varepsilon)$ show that the free energy associated with $\bar{Z}(\lambda_1,\ldots,\lambda_n)$ is given by the (multi-dimensional) Legendre–Fenchel transform of $\bar{s}(\mu_1,\ldots,\mu_n)$:

$$\bar{\phi}(\lambda_1,\ldots,\lambda_n) \equiv -\lim_{N\to\infty}\frac{1}{N}\ln\bar{Z}(\lambda_1,\ldots,\lambda_n) = \inf_{\mu_1,\ldots,\mu_n}\left[\lambda_1\mu_1 + \cdots + \lambda_n\mu_n - \bar{s}(\mu_1,\ldots,\mu_n)\right]. \tag{193}$$

We know that in general $s(\varepsilon)$ is not concave and it cannot be obtained by the Legendre–Fenchel transform of $\phi(\beta)$. We would expect the same difficulty in the inversion of (193). However, if it happens that $\bar{\phi}(\lambda_1,\ldots,\lambda_n)$ is differentiable for real λ (see Appendix C), then we are guaranteed that $\bar{s}(\mu_1,\ldots,\mu_n)$ can be obtained by the inversion of (193) and that it is therefore concave [20]

$$\bar{s}(\mu_1,\ldots,\mu_n) = \inf_{\lambda_1,\ldots,\lambda_n}\left[\lambda_1\mu_1 + \cdots + \lambda_n\mu_n - \bar{\phi}(\lambda_1,\ldots,\lambda_n)\right]. \tag{194}$$

Obviously the practical usefulness of this method resides in the fact that generally, even in the presence of phase transitions, the function $\bar{\phi}(\lambda_1,\ldots,\lambda_n)$ happens to be differentiable. In Appendix C, we present a brief proof of the differentiability of $\bar{\phi}(\lambda_1,\ldots,\lambda_n)$ and of the validity of (194) for the cases in which the global variables $\mu_k(x)$ are sums of one-particle functions.

Finally, it is not difficult to see how $\phi(\beta)$ is related to $\bar{s}(\mu_1,\ldots,\mu_n)$. Again, with steps analogous to those linking $\phi(\beta)$ to $s(\varepsilon)$, we find that $\phi(\beta)$ is given by the following extremal problem

$$\phi(\beta) = \beta f(\beta) = \inf_{\mu_1,\ldots,\mu_n}\left[\beta\bar{\varepsilon}(\mu_1,\ldots,\mu_n) - \bar{s}(\mu_1,\ldots,\mu_n)\right]. \tag{195}$$

The two variational problems (191) and (195) express the microcanonical entropy and the canonical free energy as a function of $\bar{s}(\mu_1,\ldots,\mu_n)$ and of the energy function $\bar{\varepsilon}(\mu_1,\ldots,\mu_n)$, and offer a tool to study ensemble equivalence in concrete systems.

It is important to emphasize that the large deviation method allows to reduce the statistical mechanics study of a model to an optimization problem. The method of global variables reduces exactly (or sometimes approximately) the search of the equilibrium solution to a variational problem. This approach often drastically simplifies the derivation of the statistical mechanics properties and is not very well known among physicists.

However, such a procedure does not apply to all long-range interacting systems. In particular, those for which statistical mechanics cannot be reduced (even approximately) to a mean-field variational problem are excluded. As shown, the method is strongly dependent on the possibility to introduce global or coarse-grained variables (examples are the averaged magnetization, the total kinetic energy, *etc.*) such that the Hamiltonian can be expressed as a function of these variables, as in Eq. (186), modulo a remaining term $R(x)$ whose relative contribution vanishes in the thermodynamic limit. If also short range interactions are present the procedure will not be possible. One might still be able to express the Hamiltonian as a function of coarse-grained variables plus a rest, but the rest will not vanish in the thermodynamic limit.

The global variables $\{\mu_i\}$ could be given by fields. For instance, they could correspond to a local mass density in a gravitational system, or a coarse-grained vorticity density in 2D turbulence. In these cases they would be mathematically infinite dimensional variables. However, with natural extensions, the steps that we have shown can be repeated [57,63,64].

Finally, we note that we could be interested in entropy functions depending not only on the energy density ε, but also on other quantities, see e.g. Eq. (95) for the dependence on magnetization. In this case, also these additional quantities should be expressed as functions of the global variables $\mu_1,\ldots,\mu_n$ (e.g. $m(x) = \bar{m}(\mu_1,\ldots,\mu_n)$), while the variational problem (191) will be constrained to fixed values of $\bar{\varepsilon}(\mu_1,\ldots,\mu_n) = \varepsilon$ and $\bar{m}(\mu_1,\ldots,\mu_n) = m$. Examples of this sort will be discussed in the following.

4.5.2. The three-states Potts model: An illustration of the method

Here, we apply the large deviation method to the three-state Potts model with infinite range interactions [127]. This simple example has been used as a toy model to illustrate peculiar thermodynamic properties of long-range systems [153]. The diluted three-state Potts model with short-range interactions has also been studied in connection with "Small" systems thermodynamics by Gross [154].

The Hamiltonian of the three-state Potts model is

$$H_N = -\frac{J}{2N}\sum_{i,j=1}^{N}\delta_{S_i,S_j}. \tag{196}$$

Each lattice site i is occupied by a spin variable S_i, which assumes three different states a, b, or c. A pair of spins gives a ferromagnetic contribution $-J$ ($J>0$) to the total energy if they are in the same state, and no contribution otherwise. It is important to stress that the energy sum is extended over *all* pairs (i,j): the interaction is infinite range.

Let us apply the method, following the three steps described in the following. The first step of the method consists of associating, to every microscopic configuration x, global (coarse-grained) variables, such that the Hamiltonian can be expressed as a function of them. For Hamiltonian (196) the appropriate global variables are

$$\mu \equiv (\mu_a, \mu_b) = (n_a, n_b), \tag{197}$$

where (n_a, n_b) are the fractions of spins in the two different states a, b. As a function of the microscopic configuration, we have

$$\mu_k = \frac{1}{N}\sum_{i=1}^{N}\delta_{S_i,k}, \tag{198}$$

with $k=a,b$. The Hamiltonian is expressed in terms of the global variables as

$$\bar{\varepsilon}(n_a,n_b) = -\frac{J}{2}\left[n_a^2+n_b^2+(1-n_a-n_b)^2\right], \tag{199}$$

which is exact for any N.

The second step is the computation of $\bar{s}(\mu_1,\mu_2)$. According to what we have shown in the previous subsection, we have to compute first the partition function

$$\bar{Z}(\lambda_a,\lambda_b) = \sum_{S_i}\exp\left(-\lambda_a\sum_{i=1}^{N}\delta_{S_i,a}-\lambda_b\sum_{i=1}^{N}\delta_{S_i,b}\right), \tag{200}$$

where the integral over the configurations is a discrete sum in this case. The last expression is easily solved, to find

$$\bar{Z}(\lambda_a,\lambda_b) = \left(\mathrm{e}^{-\lambda_a}+\mathrm{e}^{-\lambda_b}+1\right)^N. \tag{201}$$

We therefore obtain

$$\bar{\phi}(\lambda_a,\lambda_b) = -\ln\left(\mathrm{e}^{-\lambda_a}+\mathrm{e}^{-\lambda_b}+1\right). \tag{202}$$

As this function is evidently analytic, we can compute the function $\bar{s}$ by a Legendre–Fenchel transform. The entropy function $\bar{s}(n_a,n_b)$ will then be

$$\bar{s}(n_a,n_b) = \inf_{\lambda_a,\lambda_b}\left[\lambda_a n_a+\lambda_b n_b+\ln\left(\mathrm{e}^{-\lambda_a}+\mathrm{e}^{-\lambda_b}+1\right)\right], \tag{203}$$

which is easily solved to get

$$\bar{s}(n_a,n_b) = -n_a\ln n_a - n_b\ln n_b - (1-n_a-n_b)\ln(1-n_a-n_b). \tag{204}$$

We now proceed to the third and final step of the calculation of the entropy $s(\varepsilon)$. The variational problem (191) becomes

$$s(\varepsilon) = \sup_{n_a,n_b}\left(-n_a\ln n_a - n_b\ln n_b - (1-n_a-n_b)\ln(1-n_a-n_b)| - \frac{J}{2}\left(n_a^2+n_b^2+(1-n_a-n_b)^2\right)=\varepsilon\right). \tag{205}$$

As we have anticipated, this expression could be obtained by direct counting. The variational problem (205) can be solved numerically. The microcanonical inverse temperature can then be derived and it is shown in Fig. 22 in the allowed energy range $[-J/2,-J/6]$. Ispolatov and Cohen [153] have obtained the same result by determining the density of states. A negative specific heat region appears in the energy range $[-0.215J,-J/6]$.

Let us now consider the canonical ensemble. Applying Eq. (195) to Hamiltonian (196), we have

$$\phi(\beta) = \inf_{n_a,n_b}\left(n_a\ln n_a + n_b\ln n_b + (1-n_a-n_b)\ln(1-n_a-n_b) - \frac{\beta J}{2}\left(n_a^2+n_b^2+(1-n_a-n_b)^2\right)\right). \tag{206}$$

To obtain the caloric curve, one has to compute $\varepsilon = \mathrm{d}\phi/\mathrm{d}\beta$. Fig. 22 shows that at the canonical transition inverse temperature $\beta^t \simeq 2.75$, corresponding to the energy $\varepsilon^t/J \simeq -0.255$, a first order phase transition appears, with an associated latent heat. The low energy "magnetized" phase becomes unstable, while the high energy "homogeneous" phase, which has the constant energy density, $\varepsilon/J=-1/6$, is stabilized. In Fig. 22, the two dotted regions have the same area, respecting Maxwell's construction. At the inverse transition temperature, there is also a jump in the global variables (n_a, n_b, $1-n_a-n_b$), which are the order parameters of the model.

This extremely simple example shows already ensemble inequivalence. In the microcanonical ensemble, there is no phase transition and the specific heat becomes negative. On the other hand, in the canonical ensemble, there is a first order phase transition with a latent heat. The caloric curves do not coincide. We observe that in the energy range of ensemble inequivalence, microcanonical temperatures do not coincide with any canonical one.

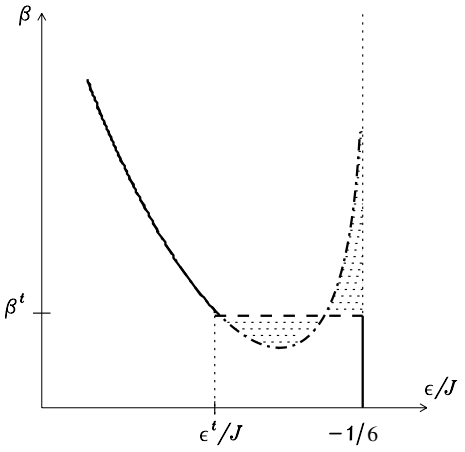


Fig. 22. Caloric curve (inverse temperature vs. energy density) for the three-states infinite range Potts model. The canonical solution is represented by a solid line. The microcanonical solution coincides with the canonical one for $\varepsilon \leq \varepsilon^t$ and is instead indicated by the dash–dotted line for $\varepsilon^t \leq \varepsilon < -J/6$. The increasing part of the microcanonical dash–dotted line corresponds to a negative specific heat region. In the canonical ensemble, the model displays a first order phase transition at β^t. The two dotted regions bounded by the dashed line and by the microcanonical dash–dotted line have the same area (Maxwell's construction).

4.5.3. The HMF model: Dealing with continuous variables

In this subsection, we consider again the HMF model to show the implementation of the large deviation method to a system with continuous variables. Let us write again here Hamiltonian (128) using the components of the magnetization, m_x and m_y, defined in (127) explicitly evidenced:

$$H_N = \sum_{i=1}^{N} \frac{p_i^2}{2} + \frac{N}{2}\left(1 - m_x^2 - m_y^2\right). \tag{207}$$

By a direct inspection of Hamiltonian (207), one can identify the global variables $u = \frac{1}{N}\sum_i p_i^2$, m_x and m_y. Since $v = \frac{1}{N}\sum_i p_i$, the average momentum, is a conserved quantity with respect to the dynamics defined by the Hamiltonian, it is convenient to included it among the global variables. Therefore, we will compute the entropy function $s(\varepsilon, v)$.

The first step of the procedure therefore consists in the identification of the following global variables

$$\mu = (u, v, m_x, m_y). \tag{208}$$

We note that also in this case the expression of the Hamiltonian as a function of the three global variables u, m_x and m_y is exact for each N, since we have

$$\bar{\varepsilon}(u, m_x, m_y) = \frac{1}{2}\left(u + 1 - m_x^2 - m_y^2\right), \tag{209}$$

without a reminder. Let us now go to the second step, i.e. the computation of $\bar{s}(u, v, m_x, m_y)$. We know, by now, that first we have to compute the partition function

$$\bar{Z}(\lambda_u, \lambda_v, \lambda_x, \lambda_y) = \int \left(\prod_i \mathrm{d}\theta_i \mathrm{d}p_i\right) \exp\left(-\lambda_u \sum_{i=1}^{N} p_i^2 - \lambda_v \sum_{i=1}^{N} p_i - \lambda_x \sum_{i=1}^{N} \cos\theta_i - \lambda_y \sum_{i=1}^{N} \sin\theta_i\right), \tag{210}$$

which is solved to get

$$\bar{Z}(\lambda_u, \lambda_v, \lambda_x, \lambda_y) = \left[\mathrm{e}^{\lambda_v^2/4\lambda_u}\sqrt{\frac{\pi}{\lambda_u}} I_0\left(\sqrt{\lambda_x^2 + \lambda_y^2}\right)\right]^N, \tag{211}$$

where I_0 is the modified Bessel function of order 0. We note that the existence of the integral in (210) requires that $\lambda_u > 0$. Then, we have

$$\bar{\phi}(\lambda_u, \lambda_v, \lambda_x, \lambda_y) = -\frac{\lambda_v^2}{4\lambda_u} - \frac{1}{2}\ln\pi + \frac{1}{2}\ln\lambda_u - \ln I_0\left(\sqrt{\lambda_x^2 + \lambda_y^2}\right). \tag{212}$$

The analyticity of this function allows us to write that

$$\bar{s}(u, v, m_x, m_y) = \inf_{\lambda_u, \lambda_v, \lambda_x, \lambda_y}\left[\lambda_u u + \lambda_v v + \lambda_x m_x + \lambda_y m_y + \frac{\lambda_v^2}{4\lambda_u} + \frac{1}{2}\ln\pi - \frac{1}{2}\ln\lambda_u + \ln I_0\left(\sqrt{\lambda_x^2 + \lambda_y^2}\right)\right]. \tag{213}$$

This variational problem can be solved explicitly, and this can be done for the "kinetic" subspace (λ_u, λ_v) separately from the "configurational" one (λ_x, λ_y), giving

$$\bar{s}(u, v, m_x, m_y) = \bar{s}_{kin}(u, v) + \bar{s}_{\rm conf}(m_x, m_y). \tag{214}$$

We note in addition that $\bar{s}_{\rm conf}(m_x, m_y)$ does not depend on m_x and m_y separately, but on the modulus $m = \sqrt{m_x^2 + m_y^2}$. This was expected, since, as already emphasized, we have degeneracy with respect to the direction of the spontaneous magnetization. Using the function $B_{\rm inv}$ previously introduced, i.e., the inverse function of I_1/I_0, we obtain

$$\bar{s}_{kin}(u, v) = \frac{1}{2} + \frac{1}{2}\ln\pi + \frac{1}{2}\ln 2(u - v^2) \tag{215}$$

$$\bar{s}_{\rm conf}(m) = -mB_{\rm inv}(m) + \ln I_0(B_{\rm inv}(m)). \tag{216}$$

Let us remark that Cauchy–Schwarz inequality implies that $u \geq v^2$.

The third step of the procedure gives the entropy function

$$s(\varepsilon, v) = \sup_{u,m}\left[\bar{s}(u, v, m)\left|\frac{u}{2} + \frac{1}{2} - \frac{m^2}{2} = \varepsilon\right.\right] \tag{217}$$

$$= \sup_{u,m}\left[\bar{s}_{kin}(u, v) + \bar{s}_{\rm conf}(m)\left|\frac{u}{2} + \frac{1}{2} - \frac{m^2}{2} = \varepsilon\right.\right] \tag{218}$$

$$= \frac{1}{2} + \frac{1}{2}\ln(4\pi) + \frac{1}{2}\ln\left(\varepsilon - \frac{1}{2} + \frac{1}{2}m^2 - \frac{1}{2}v^2\right) - mB_{\rm inv}(m) + \ln I_0(B_{\rm inv}(m)), \tag{219}$$

where in the last equality m satisfies the equation

$$\frac{m}{2\varepsilon - 1 + m^2 - v^2} - B_{\rm inv}(m) = 0. \tag{220}$$

This function depends on energy and on momentum. Maximizing with respect to v, we obtain $s(\varepsilon)$. It is easy to find that this maximum is obtained when $v = 0$ for each ε, and that the entropy $s(\varepsilon)$ is given by

$$s(\varepsilon) = \frac{1}{2} + \frac{1}{2}\ln(4\pi) + \frac{1}{2}\ln\left(\varepsilon - \frac{1}{2} + \frac{1}{2}m^2\right) - \frac{m^2}{2\varepsilon - 1 + m^2} + \ln I_0\left(\frac{m^2}{2\varepsilon - 1 + m^2}\right), \tag{221}$$

with m satisfying Eq. (220) taken at $v = 0$. We have therefore recovered the previous expressions, i.e., Eq. (151) with Eq. (152) or Eq. (185) with Eq. (182).

For completeness, we also derive the rescaled free energy $\phi(\beta) = \beta f(\beta)$. Applying Eq. (195), taking into account (209), we get

$$\phi(\beta) = \beta f(\beta) = \inf_{u,v,m}\left[\beta\left(\frac{u}{2} + \frac{1}{2} - \frac{1}{2}m^2\right) - \tilde{s}(u, v, m)\right]. \tag{222}$$

This variational problem can be easily solved to get (133).

4.5.4. A generalized HMF model: Ergodicity breaking

A generalized HMF model has been recently introduced with Hamiltonian [155,156]

$$H_N = \sum_i^N \frac{p_i^2}{2} + U, \tag{223}$$

where

$$U = NW(m) = N\left(-\frac{J}{2}m^2 - K\frac{m^4}{4}\right), \tag{224}$$

with $m = \sqrt{m_x^2 + m_y^2}$, being m_x, m_y defined in formula (127). Models of this kind have been used to describe the physics of nematic liquid crystals [157].

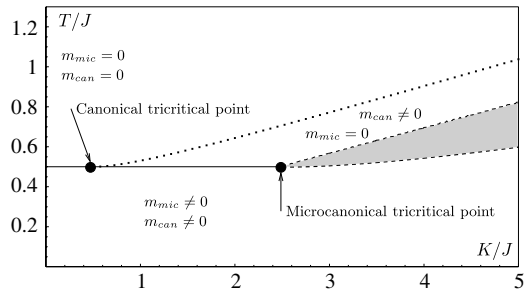


Fig. 23. Phase diagram of model (223). The canonical second order transition line (solid line at $T/J = 1/2$) becomes first order (dotted line, determined numerically) at the canonical tricritical point. The microcanonical second order transition line coincides with the canonical one below $K/J = 1/2$ but extends further right to the microcanonical tricritical point at $K/J = 5/2$. At this latter point, the transition line bifurcates in two first order microcanonical lines, corresponding to a temperature jump. The behavior of the order parameter in the two ensembles is also shown in the figure, to highlight the striking difference in the predictions of the two ensembles.

4.5.4.1. Statistical mechanics. Let us now study the statistical mechanics of model (223) with $J, K > 0$ using large deviation techniques, in a similar way as was done for the HMF model in Section 4.5.3. Like the HMF model this model has four global variables

$$\mu = (m_x, m_y, u, v) \tag{225}$$

and its microcanonical entropy is given by the same formula (217) as for the HMF model. However, the consistency equation (220) is replaced by the following

$$B_{\text{inv}}(m) = \frac{Jm + Km^3}{2\varepsilon + Jm^2 + Km^4/2 - v^2}. \tag{226}$$

One solves analytically this relation for second order phase transition and numerically for first order phase transition. Taking then into account the standard relation between inverse temperature and energy, one gets the phase diagram and the caloric curves discussed below. Similarly to the HMF model, one can also get the rescaled canonical free energy

$$\phi(\beta) = \inf_x \left[\frac{\beta J m^2}{2} + \frac{3\beta K m^4}{4} - \ln(2\pi I_0[\beta(Jm + Km^3)]) \right]. \tag{227}$$

In Fig. 23, we report the phase diagram of model (223) in the parameter space $(T/J, K/J)$ in both the canonical and the microcanonical ensemble. As expected [14], the two phase diagrams differ in the region where the canonical transition line is first order. The second order phase transition lines at $T/J = 1/2$ coincide up to the canonical tricritical point, located at $K/J = 1/2$, but then the predictions of the two ensembles differ: while the canonical ensemble gives a first order phase transition, in the microcanonical ensemble the critical line at $T/J = 1/2$ persists up to the microcanonical tricritical point at $K/J = 5/2$. Even more striking is the difference between the two ensembles for $K/J > 5/2$. Here, two different temperatures coexist at the transition energy in the microcanonical ensemble, showing what we have called a temperature jump. This corresponds to a microcanonical first order phase transition, whose "phase coexistence" region is shown by the shaded area in Fig. 23. As for the order parameter m, the most impressive difference between the two ensembles is in the intermediate region between the canonical first order phase transition line and the upper microcanonical first order line, where the canonical ensemble predicts a non-vanishing order parameter $m_{\text{can}} \neq 0$, while the microcanonical ensemble gives $m_{mic} = 0$.

Fig. 24 displays temperature vs. energy, the so-called caloric curve, in the region where both the canonical and the microcanonical ensemble predict a first order phase transition. However, the microcanonical ensemble gives two temperatures at the transition energy. Both a region of negative specific heat and a temperature jump are present in the microcanonical ensemble. The specific heat is always positive in the canonical ensemble and latent heat is released. In a microcanonical simulation performed with $N = 100$ particles (shown by the points in Fig. 24), the temperature jump is smoothed by the finite size effects and negative specific heat is well reproduced. The agreement with microcanonical thermodynamics predictions is very good.

We have shown how a simple generalization of the HMF model can give rise to all the features of ensemble inequivalence displayed by the BEG model, *i.e.* negative specific heat and temperature jumps in the microcanonical ensemble. A novel characteristic of this model with respect to the BEG model is the fact that the variables take continuous values. Moreover, the model possess a true Hamiltonian dynamics which allows us to confirm the theoretical predictions obtained in the thermodynamic limit also for finite N and to study non-equilibrium features.

4.5.4.2. Parameter space convexity. We here discuss a concrete example where we can show that the space of thermodynamic parameters is not convex, as already discussed in Section 2.3. Historically, this phenomenon has been first observed for a spin chain with asymmetric coupling [158–160], and for an Ising model with both nearest neighbor and mean-field interactions [161], see Section 6.1.1. However, it has been later realized that this occurs more generally, and even in the

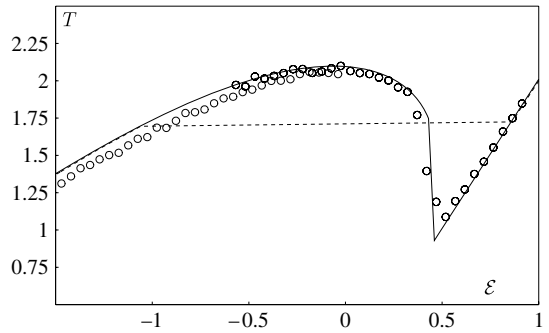


Fig. 24. Caloric curve $T(\varepsilon)$ at $K/J = 10$. The microcanonical ensemble (solid line) predicts a region of negative specific heat, where temperature T decreases as the energy is increased. Moreover, a temperature jump is present at the transition energy. In the canonical ensemble, we have a first order phase transition (dashed line). The points are the result of a molecular dynamics simulation performed by solving numerically the equations of motion given by Hamiltonian (223) with $N = 100$.

simple model studied here (223). We will consider the case $J = -1$ and $K > 0$ [156] for which the m^2 term in the Hamiltonian (223) is antiferromagnetic. We will study the structure of the set of accessible states in the space of thermodynamic parameters in the microcanonical ensemble.

Intuitively, we expect that, for large values of K, the system is ferromagnetic while, for small values of K, the antiferromagnetic coupling will dominate and makes the system paramagnetic. As we shall demonstrate below, there exists a range of values of the parameter K for which the model exhibits a first order microcanonical phase transitions between a paramagnetic phase at high energies and a ferromagnetic phase at low energies. In both phases there are regions in the (ε, K) plane in which the accessible magnetization interval exhibits a gap, resulting in breaking of ergodicity.

The specific kinetic energy $u/2 = \varepsilon - W(m)$ is by definition a non-negative quantity, which implies that

$$\varepsilon \geq W(m) = m^2/2 - Km^4/4. \tag{228}$$

We will show that as a result of this condition not all the values of the magnetization m are attainable in a certain region in the (ε, K) plane; a disconnected magnetization domain is indeed a typical case. As explained below, this situation is the one of interest. Let us characterize the accessible domains in the (ε, K) plane more precisely by analyzing the different values of K (all this will become clear after looking at Fig. 25).

For $K < 1$, the local maximum of the potential energy W is not located inside the magnetization interval $[0, 1]$ (see Fig. 25a). The potential being a strictly increasing function of the magnetization, the maximum is reached at the extremum $m = 1$. The complete interval $[0, 1]$ is thus accessible, provided the energy ε is larger than $W(1)$: the corresponding domain is in *R1* defined and illustrated in Figs. 25 and 26. The horizontally shaded region is forbidden, since the energy is lower than the minimum of the potential energy $W(0) = 0$. Finally, the intermediate region $0 < \varepsilon < W(1)$ defines region *R2*: it is important to emphasize that only the interval $[0, m_-(\varepsilon, K)]$, where $m_\pm(\varepsilon, K) = [(1 \pm \sqrt{1 - 4\varepsilon K})/K]^{1/2}$, is accessible. Larger magnetization values correspond to a potential energy $W(m)$ larger than the energy density ε, which is impossible. Fig. 25a also displays the value m_- corresponding to an energy in the intermediate region *R2*.

For $K \geq 1$, the specific potential energy W has a maximum $W_{\max} = 1/4K$ which is reached at $0 < m_{\max} = 1/\sqrt{K} \leq 1$. Fig. 25b and c, where the potential energy per particle defined in Eq. (224) is plotted vs. magnetization m, display such cases. For an energy ε larger than the critical value $W_{\max}$, condition (228) is satisfied for any value of the magnetization m. The complete interval $[0, 1]$ is thus accessible for the magnetization m. This region is *R1* represented in Fig. 26.

Let us now consider the cases for which $\varepsilon \leq W_{\max}$. As discussed above, the minimum $W_{\min}$ of the potential energy is also important to distinguish between the different regions. For $1 \leq K \leq 2$ (see Fig. 25b), the minimum of $W(m)$ corresponds to the non-magnetic phase $m = 0$ where $W(0) = 0$. The quadrilled region shown in Fig. 26, which corresponds to negative energy values, is thus not accessible. On the contrary, positive energy values are possible and correspond to very interesting cases, since only sub-intervals of the complete magnetization interval $[0, 1]$ are accessible. There are however two different cases

- for $0 < W(1) < \varepsilon < W_{\max}$, the domain of possible magnetizations is $[0, m_-(\varepsilon, K)] \cup [m_+(\varepsilon, K), 1]$. The above conditions are satisfied in *R3* of Fig. 26.
- for $0 < \varepsilon < W(1)$, only the interval $[0, m_-(\varepsilon, K)]$ satisfies condition (228). This takes place within *R2* of Fig. 26.

For the domain $2 \leq K$, the minimum of the potential energy is attained at the extremum, $m = 1$, implying $\varepsilon > W(1) = 1/2 - K/4$. The vertically shaded region is thus forbidden. In the accessible region, two cases can be identified

- for $W(1) < \varepsilon < 0$, only the interval $[m_+(\varepsilon, K), 1]$ satisfies condition (228). It is important to note that $m_+(\varepsilon, K) \leq 1$ provided $\varepsilon \geq 1/2 - K/4$. These cases correspond to region *R4*.
- for $0 \leq \varepsilon \leq W_{\max}$, the two intervals $[0, m_-(\varepsilon, K)]$ and $[m_+(\varepsilon, K), 1]$ satisfy condition (228), corresponding to *R3* of Fig. 26.

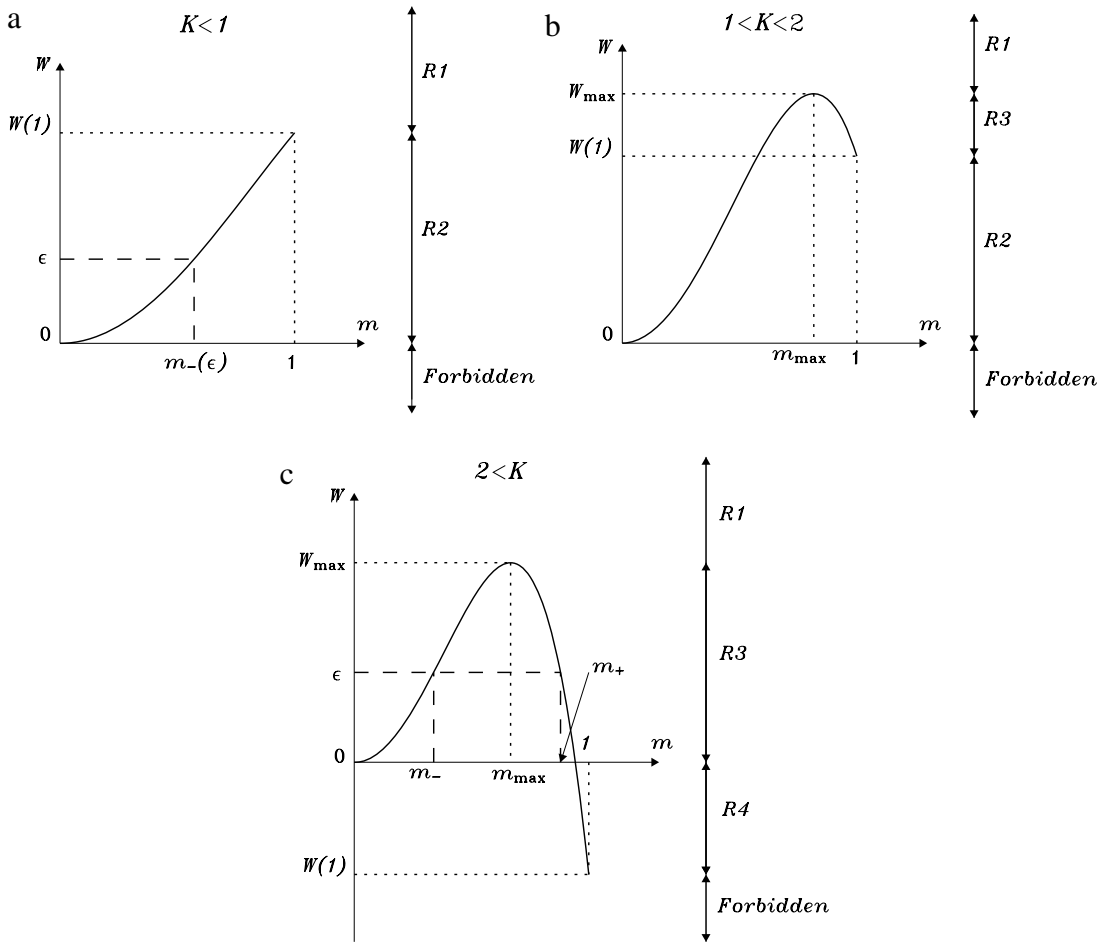


Fig. 25. Specific potential energy W vs. magnetization m for three different cases: $K < 1$ (panel a), $1 < K < 2$ (panel b) and $2 < K$ (panel c). The location of the maximal magnetization $m_{\max}$ and the corresponding potential energy $W_{\max}$ are shown (see text). In panels (a) and (c), two examples of the location of the critical magnetization $m_{\pm}(\varepsilon, K)$ are indicated for energy density values ε in the intermediate regions.

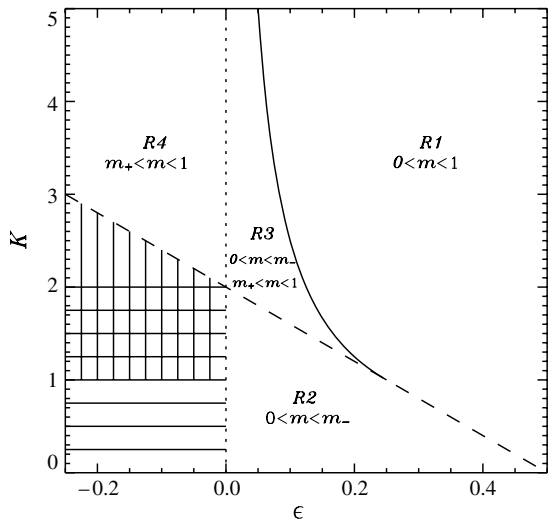


Fig. 26. The (ε, K) plane is divided in several regions. The solid curve corresponds to $K = 1/(4\varepsilon)$, the oblique dashed line to $K = 2 - 4\varepsilon$, while the dotted one to $K = 1$. The vertically shaded, quadrilled, and horizontally shaded regions are forbidden. The accessible magnetization interval in each of the four regions is indicated (see text for details).

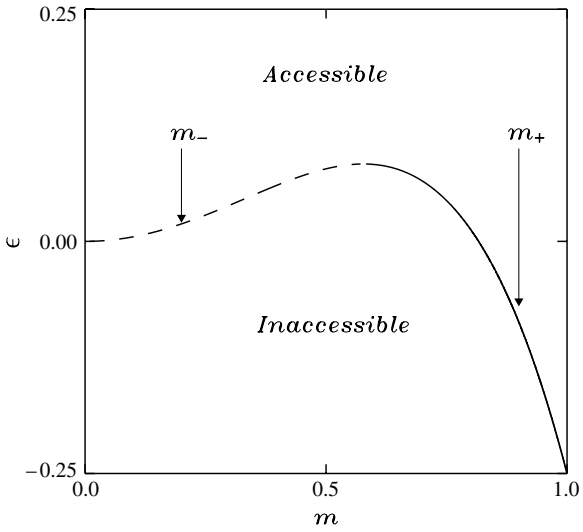


Fig. 27. Accessible region in the (m, ε) plane for $K = 3$. For energies in a certain range, a gap in the accessible magnetization values is present and defined by the two boundaries $m_\pm(\varepsilon, K)$.

In summary, the complete magnetization interval $[0, 1]$ is accessible only in the region *R1*. In *R2*, only $[0, m_-]$ is accessible, while only $[m_+, 1]$ is accessible in *R4*. Finally, we note that the phase space of the system is not connected in the region *R3*. Indeed, the magnetization cannot vary continuously from the first interval $[0, m_-]$ to the second one $[m_+, 1]$, although both are accessible. These restrictions yield the accessible magnetization domain shown in Fig. 27. The fact that for a given energy the space of the thermodynamic parameter m is disconnected implies ergodicity breaking for the Hamiltonian dynamics. It is important to emphasize that the discussion above is independent of the number of particles and ergodicity is expected to be broken even for a finite N.

4.5.4.3. Phase diagram in the microcanonical ensemble. We have thus found that in certain regions in the (ε, K) plane, the magnetization cannot assume any value in the interval $[0, 1]$. For a given energy there exists a gap in this interval to which no microscopic configuration can be associated. Let us now study the phase diagram in the microcanonical ensemble, i.e. in the parameter space (ε, K).

First, by comparing the low and high energy regimes, it is possible to show that a phase transition is present between the two regimes. In the domain *R4* of Fig. 26, for very low energy ε (close to the limiting value $1/2 - K/4$), the accessible range for m is a small interval located close to $m = 1$ (see Fig. 25c). The maximum in m of the entropy $\tilde{s}(\varepsilon, m)$ corresponds therefore to a magnetized state located very close to $m = 1$ (see the top-left inset in Fig. 28). On the contrary, in the very large energy domain ($\varepsilon \gg m \simeq 1$), the variations of the entropy with respect to m are dominated by the variations of the configurational entropy $\tilde{s}_{\text{conf}}$, since the kinetic entropy

$$\tilde{s}_{kin}(\varepsilon, m) = \frac{1}{2} \ln \left(\varepsilon - \frac{m^2}{2} + K \frac{m^4}{4} \right) \tag{229}$$

is roughly a constant $\tilde{s} \simeq (\ln \varepsilon)/2$ when m is of order one. As expected, the configurational entropy is a decreasing function of the magnetization: the number of microstates corresponding to a paramagnetic macrostate being much larger than the same number for a ferromagnetic state. The configurational entropy has therefore a single maximum located at $m^* = 0$. A phase transition takes place between the paramagnetic state at large energy and a magnetic state at small energy. Moreover, as the paramagnetic state is possible only for positive energies ε (see Fig. 25), the transition line is located in the domain $\varepsilon \geq 0$. In this region, the quantity $\partial_m^2 \tilde{s}(\varepsilon, 0) = -(1 + 1/(2\varepsilon))$ is negative, which ensures that, for any value of ε and K, the paramagnetic state $m^* = 0$ is a local entropy maximum. The latter argument allows us to exclude a second order phase transition at a positive critical energy, since the second derivative $\partial_m^2 \tilde{s}(\varepsilon, 0)$ would have to vanish, which is impossible. The above argument leads to the conclusion that the phase transition must be *first order*.

Let us now focus on the behavior of the entropy in the region *R3*, where the accessible range for m is the union of two disconnected intervals $[0, m_-] \cup [m_+, 1]$. As discussed above, the entropy $\tilde{s}(\varepsilon, m)$ has a local maximum in the first interval $[0, m_-]$ located at $m^* = 0$ and associated with the entropy $\tilde{s}^1_{\max} = \tilde{s}(\varepsilon, 0) = \log(\varepsilon)/2$. In the second interval $[m_+, 1]$, a maximum is also present with $\tilde{s}^2_{\max} = \tilde{s}(\varepsilon, m^*)$ where $m^* \geq m_+ > 0$. As $\tilde{s}^1_{\max}(\varepsilon)$ diverges to $-\infty$ when ε tends to 0, a magnetized state is expected on the line $\varepsilon = 0$, as long as $\tilde{s}^2_{\max}$ remains finite. Since $K = 2$ is the only value for which $\tilde{s}^2_{\max}(0, K)$ diverges, the first order transition line originates at the point $B(0, 2)$ in Fig. 28. Although it is possible to study analytically

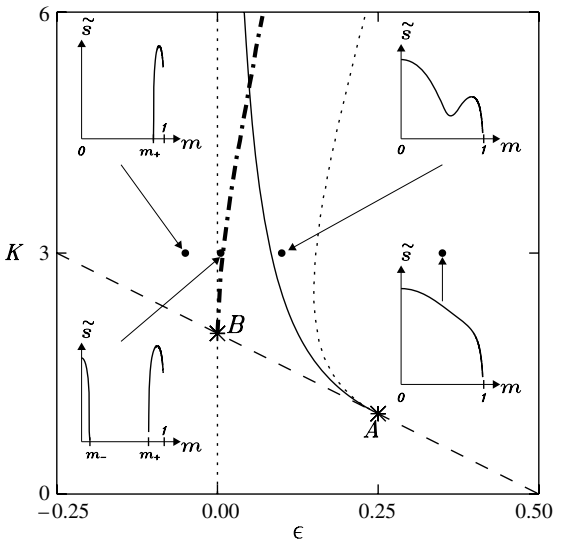


Fig. 28. Phase diagram of the mean field model (223) with $J = -1$. The dash–dotted curve corresponds to the first order phase transition line, issued from the point $B(0, 2)$. As in Fig. 26, the solid curve indicates the right border of the region *R3*, where the space of thermodynamic parameters is disconnected. The dashed line corresponds to $K = 2-4\varepsilon$. The dotted line issued from the point $A(1/4,\ 1)$ represents the metastability line for the magnetized state, while the $\varepsilon = 0$ vertical dotted line is also the metastability line for the paramagnetic state. The four insets represent the entropy $\tilde{s}$ versus the magnetization m for the four energies: $\varepsilon = -0.05$, 0.005, 0.1 and 0.35, when $K = 3$.

the asymptotic behavior of the transition line near this point, we can rather easily compute numerically the location of the first order transition line, represented by the dash–dotted line in Fig. 28.

Fig. 28 also shows the metastability line (the dotted line starting at point $A(1/4,\ 1)$), for the magnetized state $m^* \neq 0$. To the right of this metastability line, there is no metastable state (local entropy maximum for any $m > 0$, see bottom-right inset in Fig. 28) while a metastable state (local maximum) exists at some non-vanishing magnetization on the other side (see top-right inset in Fig. 28). Finally, the vertical dotted line $\varepsilon = 0$ corresponds to the metastability line of the paramagnetic state $m^* = 0$.

One of the key issues we would like to address is the possible links between the breakdown of connectivity, and thus ergodicity breaking, on the one hand, and the phase transition, on the other hand. Obvious general properties do exist: a region of parameters where the space is disconnected corresponds to a region where metastable states exist. Let us justify this statement. At the boundary of any connected domain, when the order parameter is close to its boundary value m_b, there is a single accessible state. In a model with continuous variables, like the one we discuss here, this leads to a divergence of the entropy. In this case the singularity of the entropy is proportional to $\ln(m - m_b)$ (see for instance Eq. (229)). For a model with discrete variables, like an Ising model, the entropy no longer reaches $-\infty$ as m tends to m_b, but would rather take a finite value. However, the singularity would still exist and would then be proportional to $(m - m_b)\ln(m - m_b)$. In both cases, of discrete and continuous variables, at the boundary of any connected domain, the derivative of the entropy as a function of the order parameter tends generally to $\pm\infty$. As a consequence, entropy extrema cannot be located at the boundary. Thus, a local entropy maximum (metastable or stable) does exist in a region of parameters where the space is disconnected.

Hence, there is an entropy maximum (either local or global) in any connected domain of the space. For instance, considering the present model, in Fig. 28, the area *R3* is included in the area where metastable states exist (bounded by the two dotted lines and the dashed line). In such areas where metastable states exist, one generically expects first order phase transitions. Thus the breakdown of phase space connectivity is generically associated to first order phase transition, as exemplified by the present study. However, this is not necessary, one may observe metastable states without first order phase transitions, or first order phase transitions without connectivity breaking.

A very interesting question is related to the critical points *A* and *B* shown in Fig. 28. As observed in the phase diagram, the end point for the line of first order phase transition (point *B*) corresponds also to a point where the boundary of the region where the space is disconnected is not smooth. Similarly, the end point for the line of appearance of metastable states (point *A*) is also a singular point for the boundary of the area where the space is disconnected. It is thus possible to propose the conjecture that such a relation is generic, and that it should be observed in other systems where both first order phase transitions and phase space ergodicity breaking do occur.

4.5.4.4. Equilibrium dynamics. The feature of disconnected accessible magnetization intervals, which is typical of systems with long-range interactions, has profound implications on the dynamics. In particular, starting from an initial condition which lies within one of these intervals, local dynamics is unable to move the system to a different accessible interval. A macroscopic change would be necessary to carry the system from one interval to the other. Thus the ergodicity is broken in the microcanonical dynamics even at finite N.

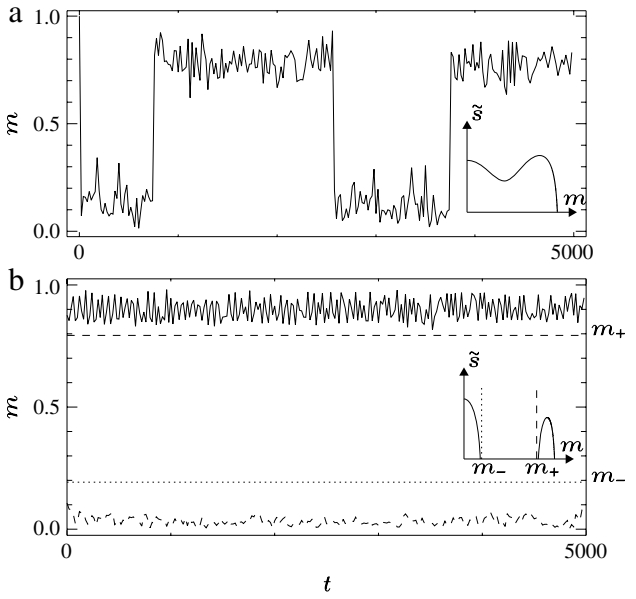


Fig. 29. Time evolution of the magnetization m (the entropy of the corresponding cases is plotted as an inset). Panel (a) corresponds to the case $\varepsilon = 0.1$ and $K = 8$, while panel (b) to $\varepsilon = 0.0177$ and $K = 3$. In panel (b), two different initial conditions are plotted simultaneously: the solid line corresponds to $m(t = 0) = 0.1$ while the dashed line to $m(t = 0) = 0.98$. The dashed (resp. dotted) line in panel (b) corresponds to the line $m = m_+ \simeq 0.794$ (resp. $m = m_- \simeq 0.192$).

In Ref. [161], this point has been demonstrated using the microcanonical Monte-Carlo dynamics suggested by Creutz [162], see Section 6.1.1. Here, we use the Hamiltonian dynamics given by the equations of motion

$$\dot{\theta}_n = +\frac{\partial H_N}{\partial p_n} = p_n \tag{230}$$

$$\dot{p}_n = -\frac{\partial H}{\partial \theta_n} = N\left(1 - Km^2\right)\left(\sin\theta_n m_x - \cos\theta_n m_y\right). \tag{231}$$

We display in Fig. 29 the evolution of the magnetization for two cases, since we have shown above that the gap opens up when ε decreases. The first case corresponds to the domain *R1*, in which the accessible magnetization domain is the full interval [0, 1]. Fig. 29a presents the time evolution of m. The magnetization switches between the paramagnetic metastable state $m^* = 0$ and the ferromagnetic stable one $m^* > 0$. This is possible because the number of particles is small ($N = 20$) and, as a consequence, the entropy barrier (see the inset) can be overcome. Considering a system with a small number of particles allows to observe flips between local maxima, while such flips would be less frequent for larger N values.

In the other case, we consider a stable $m^* = 0$ state which is disconnected from the metastable one. This makes the system unable to switch from one state to the other. Note that this feature is characteristic of the microcanonical dynamics, since an algorithm reproducing canonical dynamics would allow the crossing of the forbidden region (by moving to higher energy states, which is impossible in the microcanonical ensemble). The result of two different numerical simulations is reported in Fig. 29b. One is initialized with a magnetization within $[0, m_-]$, while the other corresponds to an initial magnetization close to $m(0) = 1$ (*i.e.* within $[m_+, 1]$). One clearly sees that the dynamics is blocked in one of the two possible regions, and not a single jump is visible over a long time span. This is a clear evidence of ergodicity breaking.

4.5.5. The mean-field ϕ^4 spin model: Negative susceptibility

We here show how to solve the so called mean-field ϕ^4 spin model using large deviations. The Hamiltonian of the model is the following

$$H_N = \sum_{i=1}^{N}\left(\frac{p_i^2}{2} - \frac{1}{4}q_i^2 + \frac{1}{4}q_i^4\right) - \frac{1}{4N}\sum_{i,j=1}^{N} q_i q_j. \tag{232}$$

It is a system of unit mass particles moving on a line. These particles are subjected to a local double-well potential, and they interact with each other through a mean-field (infinite range) interaction given by the all-to-all coupling in the double sum.

This model was introduced by Desai and Zwanzig [163]. More recently, the canonical solution was obtained [164], showing that the system exhibits a second-order ferromagnetic phase transition at a critical temperature $T_c \simeq 0.264$,

corresponding to a critical energy $\varepsilon_c = T_c/2 \simeq 0.132$. Later, the entropy of the ϕ^4 model in presence of an extra magnetic field was computed [165]. Finally, a calculation of the entropy, as a function of energy and magnetization

$$m = \frac{1}{N}\sum_{i=1}^{N} q_i, \tag{233}$$

has been performed in Refs. [116,166,167]. We note that in this model the modulus of this quantity is not bounded inside the interval $[-1, 1]$.

We will perform the three steps of the large deviation method, in order to compute the microcanonical entropy as a function of the energy ε and of the magnetization m; i.e., we will compute the function $\tilde{s}(\varepsilon, m)$ defined in Eq. (95). More details can be found in Ref. [116].

The first step consists of the identification of the global variables in terms of which we can express the energy ε. In addition to magnetization other global variables are u, twice the average kinetic energy, and

$$z = \frac{1}{4N}\sum_{i=1}^{N}\left(q_i^4 - q_i^2\right), \tag{234}$$

related to the local potential. We easily see that

$$\bar{\varepsilon}(u, z, m) = \frac{1}{2}u + z - \frac{1}{4}m^2. \tag{235}$$

The second step begins with the computation of the partition function

$$\bar{Z}(\lambda_u, \lambda_z, \lambda_m) = \int\left(\prod_i \mathrm{d}q_i\mathrm{d}p_i\right)\exp\left[-\lambda_u\sum_{i=1}^{N}p_i^2 - \lambda_m\sum_{i=1}^{N}q_i - \lambda_z\sum_{i=1}^{N}\left(q_i^4 - q_i^2\right)\right]. \tag{236}$$

Also in this case the calculation splits into the one of the kinetic part and of the potential part, and we have

$$\bar{Z}(\lambda_u, \lambda_z, \lambda_m) = \bar{Z}_u(\lambda_u)\bar{Z}_{z,m}(\lambda_z, \lambda_m) = \left[\sqrt{\frac{\pi}{\lambda_u}}\right]^{\frac{N}{2}}\left[\int_{-\infty}^{+\infty}\mathrm{d}q\mathrm{e}^{-\lambda_m q - \lambda_z\left(q^4 - q^2\right)}\right]^N, \tag{237}$$

where the two terms on the right-hand side define $\bar{Z}_u(\lambda_u)$ and $\bar{Z}_{z,m}(\lambda_z, \lambda_m)$, respectively. The existence of the integral requires $\lambda_u > 0$ and $\lambda_z > 0$. From the previous expression, we find

$$\bar{\phi}(\lambda_u, \lambda_z, \lambda_m) = \bar{\phi}_u(\lambda_u) + \bar{\phi}_{z,m}(\lambda_z, \lambda_m), \tag{238}$$

with

$$\bar{\phi}_u(\lambda_u) = -\frac{1}{2}\ln\pi + \frac{1}{2}\ln\lambda_u \tag{239}$$

$$\bar{\phi}_{z,m}(\lambda_z, \lambda_m) = -\ln\left[\int_{-\infty}^{+\infty}\mathrm{d}q\mathrm{e}^{-\lambda_m q - \lambda_z\left(q^4 - q^2\right)}\right]. \tag{240}$$

Both these functions are analytic, and therefore we can write

$$\bar{s}(u, z, m) = \bar{s}_{kin}(u) + \bar{s}_{\mathrm{conf}}(z, m), \tag{241}$$

where

$$\bar{s}_{kin}(u) = \inf_{\lambda_u}\left[\lambda_u u - \bar{\phi}_u(\lambda_u)\right] = \frac{1}{2} + \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln u \tag{242}$$

$$\bar{s}_{\mathrm{conf}}(z, m) = \inf_{\lambda_z, \lambda_m}\left[\lambda_z z + \lambda_m m - \bar{\phi}_{z,m}(\lambda_z, \lambda_m)\right]. \tag{243}$$

The third step of the procedure gives the entropy function

$$\tilde{s}(\varepsilon, m) = \sup_{u,z}\left[\bar{s}_{kin}(u) + \bar{s}_{\mathrm{conf}}(z, m)\,\middle|\,\frac{u}{2} + z - \frac{m^2}{4} = \varepsilon\right], \tag{244}$$

that can be rewritten as

$$\tilde{s}(\varepsilon, m) = \frac{1}{2} + \frac{1}{2}\ln(4\pi) + \sup_z\left[\frac{1}{2}\ln\left(\varepsilon - z + \frac{m^2}{4}\right) + \bar{s}_{\mathrm{conf}}(z, m)\right]. \tag{245}$$

This variational problem can be solved numerically and the result is plotted in Fig. 30 for different values of ε.

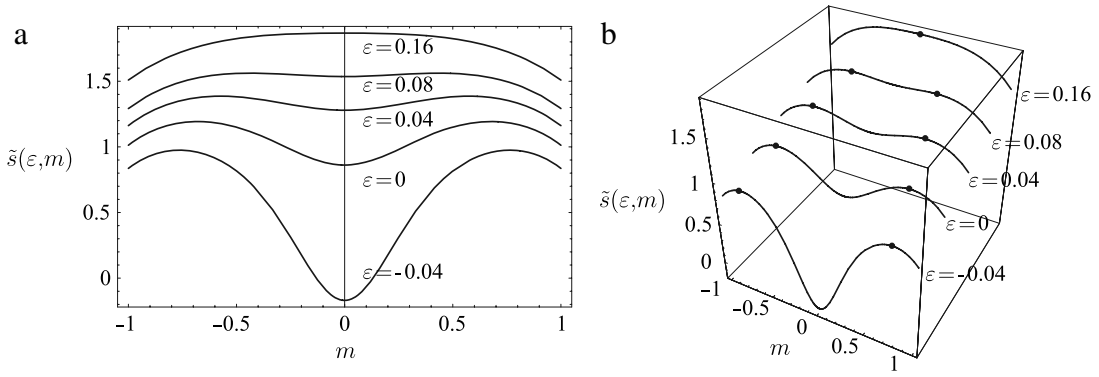


Fig. 30. (a) Entropy $\tilde{s}$ as a function of magnetization m for different values of the energy ε. (b) 3D view of $\tilde{s}(\varepsilon, m)$. The black dots show the location of the equilibrium values m^* in the microcanonical ensemble in which the magnetization constraint is released, and therefore the equilibrium at a given ε is realized for the m^* value maximizing $\tilde{s}(\varepsilon, m)$.

The purpose of this subsection, besides the possibility to offer a further implementation of the large deviation method, is to show an example of ensemble inequivalence which is not related to the presence of a negative specific heat in the microcanonical ensemble, but rather to the negativity of another quantity that in the canonical ensemble is positive definite, namely the magnetic susceptibility at constant temperature (see Section 4.3.3). Although striking, this feature appears for other long-range systems whose entropy depends on two macroscopic variables. For instance, it has been recently discussed in connection with models of two-dimensional geophysical flows [68], for which entropy depends on total circulation and energy.

The inequivalence has to be considered with respect to the canonical ensemble, whose partition function is

$$Z(\beta, h) = \int \left(\prod_i \mathrm{d}q_i \mathrm{d}p_i\right) \exp\left(-\beta\left[\sum_{i=1}^{N}\left(\frac{p_i^2}{2} - \frac{q_i^2}{4} + \frac{q_i^4}{4}\right) - \frac{1}{4N}\sum_{i,j=1}^{N} q_i q_j - h\sum_{i=1}^{N} q_i\right]\right). \tag{246}$$

From this we derive the rescaled free energy

$$\phi(\beta, h) = \beta f(\beta, h) = -\lim_{N\to\infty}\frac{1}{N}\ln Z(\beta, h) = \inf_{\varepsilon, m}\left[\beta\varepsilon - \beta h m - \tilde{s}(\varepsilon, m)\right]. \tag{247}$$

It is straightforward to apply to this case expression (195), obtaining the free energy in the framework of the large deviation method. In the concrete case of the ϕ^4 model we have

$$\beta f(\beta, h) = \inf_{u,z,m}\left[\beta\bar{\varepsilon}(u, z, m) - \beta h m - \bar{s}(u, z, m)\right]. \tag{248}$$

Our presentation in Section 4.3.3 already contains all we need to study magnetic susceptibility. We repeat here the basic formula in Eq. (125) and apply it to the ϕ^4 model.

$$\chi = \left(\frac{\partial m}{\partial h}\right) = -\beta\frac{\tilde{s}_{\varepsilon\varepsilon}}{\tilde{s}_{\varepsilon\varepsilon}\tilde{s}_{mm} - \tilde{s}_{\varepsilon m}^2}, \tag{249}$$

which is valid for both the canonical and the microcanonical ensemble. It turns out that $\tilde{s}_{\varepsilon\varepsilon}$ is always negative, and therefore the sign of the susceptibility is related to the sign of $\tilde{s}_{mm}$. We see that, whenever this quantity is positive (something that cannot happen in the canonical ensemble), the susceptibility is negative, and, by continuity, it will be negative also in regions where s_{mm} is negative.

In the left picture of Fig. 30 we plot $\tilde{s}(\varepsilon, m)$ as a function of m for different values of ε, clearly showing the regions of convexity of $\tilde{s}$. The right picture offers a 3D view of $\tilde{s}(\varepsilon, m)$. The plot of $\tilde{s}(\varepsilon, m)$ versus m has a bimodal shape, and a region of convexity, for $\varepsilon < \varepsilon_c$.

Just for the sake of clarity and readability, we can repeat an argument analogous to the one presented in Section 4.1.3, but now with the magnetization m playing the role there played by the energy ε. In that case a region of convexity in the microcanonical $s(\varepsilon)$ was associated to the presence of a first order phase transition in the canonical ensemble, with an energy jump at a given temperature, without any phase transition in the microcanonical ensemble. Now, a region of convexity in $\tilde{s}(\varepsilon, m)$, seen as a function of m, is associated again with a first order transition in the canonical ensemble. The transition is present at temperatures lower than the critical temperature T_c, and it is expressed by the jump from a positive equilibrium magnetization to a negative equilibrium magnetization, when the magnetic field h is varied from 0^+ to 0^-. Since the Hamiltonian of this model has a symmetry for $q_i \to -q_i$, the critical magnetic field, in the canonical ensemble, is $h_c = 0$. Again, no phase transition is present in the microcanonical ensemble. The analogy between the two cases is not surprising, since they are both related to the non-concavity of the entropy function, as seen with respect to two different variables.

4.5.6. The Colson–Bonifacio model for the Free Electron Laser

In this subsection we discuss an extremely simplified model of the Free Electron Laser (FEL) [168,169], which can be explicitly solved using large deviations [126,127]. It should be remarked that a model of this kind was historically introduced by Zaslavsky and coworkers [170] to describe structural phase transitions in crystals.

In the linear FEL, a relativistic electron beam propagates through a spatially periodic magnetic field, interacting with the co-propagating electromagnetic wave. Lasing occurs when the electrons bunch in a subluminar beat wave [171]. Scaling away the time dependence of the phenomenon and introducing appropriate variables, it is possible to catch the essence of the asymptotic state by studying the classical Hamiltonian

$$H_N = \sum_{j=1}^{N} \frac{p_j^2}{2} - N\delta A^2 + 2A \sum_{j=1}^{N} \sin(\theta_j - \varphi). \tag{250}$$

The p_i's are related to the energies relative to the center of mass of the N electrons and the conjugated variables θ_i characterize their positions with respect to the co-propagating wave. The complex electromagnetic field variable, $\mathbf{A} = A\,\mathrm{e}^{\mathrm{i}\varphi}$, defines the amplitude and the phase of the dominating mode ($\mathbf{A}$ and $\mathbf{A}^\star$ are conjugate variables). The parameter δ measures the average deviation from the resonance condition. In addition to the energy H_N, the total momentum $P = \sum_j p_j + NA^2$ is also a conserved quantity. Most of the studies of this model have concentrated on the numerical solution of Hamiltonian (250), starting from initial states with a small field amplitude A and the electrons uniformly distributed with a small kinetic energy. Then, the growth of the field has been observed and its asymptotic value determined from the numerics. Our study below allows to find the asymptotic value of the field analytically.

In the first step, similarly to the HMF case, Hamiltonian (250) can be rewritten as

$$H_N \simeq NH(\mu) = N\left(\frac{u}{2} - \delta A^2 + 2A\left(-m_x \sin\varphi + m_y \cos\varphi\right)\right) \tag{251}$$

where m_x, m_y, are defined in Eq. (127) and u just below Eq. (207), together with v. Defining the phase of the mean field φ' as $m_x + \mathrm{i}m_y = m\exp(\mathrm{i}\varphi')$, the global variables are $\mu = (u, v, m, \varphi', A, \varphi)$.

Going to step two we should first remark that the two field variables A, φ, give a contribution of order $1/N$ to the entropy, which can be neglected. Hence, the $\bar{\phi}$ function reduces to the one of the HMF model, see formula (212). Analogously, one obtains the same contributions to the kinetic and configurational entropies of the HMF model, shown in formulas (215) and (216).

Performing finally the third step, after defining the total momentum density as $\sigma = P/N$, the microcanonical variational problem to be solved is

$$s(\varepsilon, \sigma, \delta) = \sup_{\mu}\left[\frac{1}{2}\ln(u - v^2) + \tilde{s}_{\mathrm{conf}}(m)\,\middle|\, \varepsilon = \frac{u}{2} + 2Am\sin\left(\varphi' - \varphi\right) - \delta A^2,\ \sigma = v + A^2\right]. \tag{252}$$

Using the constraints of the variational problem, one can express u and v as functions of the other variables, obtaining the following form of the entropy

$$s(\varepsilon, \sigma, \delta) = \sup_{A,\varphi,m,\varphi'}\left[\frac{1}{2}\ln\left[2\left(\varepsilon - \frac{\sigma^2}{2}\right) - 4Am\sin\left(\varphi' - \varphi\right) + 2(\delta - \sigma)A^2 - A^4\right] + s_{\mathrm{conf}}(m)\right]. \tag{253}$$

The extremalization over the variables φ and φ' is straightforward, since by direct inspection of formula (253), it is clear that the entropy is maximized when $\varphi' - \varphi = -\pi/2$. Then

$$s(\varepsilon, \sigma, \delta) = \sup_{A,m}\left[\frac{1}{2}\ln\left[2\left(\varepsilon - \frac{\sigma^2}{2}\right) + 4Am + 2(\delta - \sigma)A^2 - A^4\right] + \tilde{s}_{\mathrm{conf}}(m)\right] \equiv \sup_{A,m}\tilde{s}(A, m). \tag{254}$$

The non-zero σ case can be reduced to the vanishing σ problem using the identity $s(\varepsilon, \sigma, \delta) = s(\varepsilon - \sigma^2/2, 0, \delta - \sigma)$. From now on, we will discuss only the zero momentum case since we can absorb non-zero momenta into the definition of ε and δ. This has also a practical interest, because it is the experimentally relevant initial condition.

The conditions for having a local stationary point are

$$\frac{\partial\tilde{s}}{\partial A} = \frac{2\left(\delta A - A^3 + m\right)}{2\varepsilon + 2\delta A^2 + 4Am - A^4} = 0, \tag{255}$$

$$\frac{\partial\tilde{s}}{\partial m} = \frac{2A}{2\varepsilon + 2\delta A^2 + 4Am - A^4} - B_{\mathrm{inv}}(m) = 0, \tag{256}$$

where B_{inv} is defined in formula (220). It is clear that $m = A = 0$ is a solution of conditions (255) and (256): it exists only for positive ε. We will limit ourselves to study its stability. It must be remarked that this is the typical initial condition studied experimentally in the FEL: it corresponds to having a beat wave with zero amplitude and the electrons uniformly distributed. The lasing phenomenon is revealed by an exponential growth of both A and the electron bunching parameter m.

The second order derivatives of the entropy $\tilde{s}(A, m)$, computed on this solution, are

$$\frac{\partial^2\tilde{s}}{\partial A^2}(0,0) = \frac{\delta}{e}, \qquad \frac{\partial^2\tilde{s}}{\partial m^2}(0,0) = -2, \qquad \frac{\partial^2\tilde{s}}{\partial A\partial m}(0,0) = \frac{1}{\varepsilon}. \tag{257}$$

The two eigenvalues of the Hessian are the solutions of the equation

$$x^2 - x\left(-2 + \frac{\delta}{\varepsilon}\right) - \frac{2\delta}{\varepsilon} - \frac{1}{\varepsilon^2} = 0. \tag{258}$$

The stationary point is a maximum if the roots of this equation are both negative. This implies that their sum $(-2 + \delta/\varepsilon)$ is negative and their product $(-2\delta/\varepsilon - 1/\varepsilon^2)$ is positive. Recalling that we restrict to positive ε values, the condition for the sum to be negative is $\varepsilon > \delta/2$ and the one for the product to be positive is $\varepsilon > -1/(2\delta)$ with $\delta < 0$. The second condition is more restrictive, hence the only region where the solution $m = A = 0$ exists and is stable is $\varepsilon > -1/(2\delta)$ with $\delta < 0$. When crossing the line $\varepsilon = -1/(2\delta)$ $(\delta < 0)$, a non-zero bunching solution $(m \neq 0)$ originates continuously from the zero bunching one, producing a second order phase transition. This analysis fully coincides with the one performed in the canonical ensemble in Refs. [172,173]. The maximum entropy solution in the region complementary to the one where the zero bunching solution is stable can be obtained by solving numerically Eqs. (255) and (256). This corresponds to having a non-zero field intensity and bunching.

The equilibrium solution of the Colson–Bonifacio model has been shown to describe well the system only in certain energy ranges [174]. Dynamical effects must be carefully taken into account. The dynamics can be studied in the $N \to \infty$ limit by a Vlasov equation, see Section 5. The homogeneous zero field state is linearly Vlasov stable for $\delta > \delta_c = (27/4)^{1/3}$ [175], hence, in this region of parameters one cannot expect that the system will relax to the equilibrium state. Moreover, as we will discuss, Vlasov equation can display an infinity of quasi-stationary states, and the system can be trapped in such states for a time that diverges with N. The value of bunching parameter and field intensity can differ from the equilibrium ones significantly [126].

4.6. The origin of singularities of thermodynamic functions

We have seen that phase transitions have a paramount importance in the study of ensemble equivalence and inequivalence. Indeed, we have emphasized that inequivalence occurs when phase transitions have different properties in the microcanonical and in the canonical ensemble. The BEG model has shown many aspects of this (see Section 4.2): the microcanonical transition lines do not coincide with the canonical ones; there is a region in the space of thermodynamic parameters where there is a first order phase transition in the canonical ensemble while the transition is second order in the microcanonical ensemble, which moreover presents a negative specific heat. There is another region where both ensembles have a first order phase transition, but with different properties, i.e. an energy jump in the canonical ensemble and a temperature jump (with sometimes also a negative specific heat) in the microcanonical ensemble. The ϕ^4 model has shown another kind of inequivalence: instead of being associated to the thermodynamically conjugated pair (ε, T), we have found that it is related to the pair (m, h).

Obviously, phase transitions are present also in short-range systems, but in that case ensembles are equivalent. We remind that, as emphasized in Section 4.1.2, we adopt here the physical point of view about ensemble equivalence at first-order phase transitions for short-range systems. At variance with the mathematical physics literature [99], we do not use the term partial equivalence in cases where there is not a one-to-one correspondence between energy in the microcanonical ensemble and temperature in the canonical ensemble.

Then, the questions may arise about a possible qualitative difference between phase transitions in short-range and in long-range systems. This section is devoted to a brief discussion of some results that concern this problem. We will give only few details, for two reasons: first, this subject lies outside the purpose of this review; second, two recent reviews and a book have appeared dealing with this problem where the interested reader can find more details [176–178]. We will discuss the origin of the analytical properties of entropy and of free energy that, in turn, are associated to the presence of phase transitions.

The ϕ^4 model can be studied without the magnetization constraint in the microcanonical ensemble, and without the presence of an external magnetic field in the canonical ensemble. This is what has been done in Ref. [164], where, as we have mentioned in Section 4.5.5, a second order phase transition has been found. According to what we have emphasized at the end of Section 4.1.3, a second order phase transition implies that the two ensembles are equivalent. Nevertheless, the presence of the transition means that the second derivatives of $s(\varepsilon)$ and of $\phi(\beta)$ are discontinuous, the first at the transition energy ε_c and the second at the corresponding transition inverse temperature β_c. The non-analyticity of the free energy is the usual signature of phase transitions. What we are interested in here is the origin of these non-analyticities.

The entropy function $\tilde{s}(\varepsilon, m)$ of the ϕ^4 model, previously studied, is an analytic function in both arguments. We have also seen that the inequivalence with the canonical ensemble, where a first order phase transition is present, is due to the concavity properties of $\tilde{s}(\varepsilon, m)$. If we study the system without the magnetization constraint, we can use, as before, the large deviation method, with the only difference that at the third step of the procedure, we maximize in Eq. (245) also with respect to m. This is simply because the entropy $s(\varepsilon)$ is, for each ε, the maximum over m of $\tilde{s}(\varepsilon, m)$. Since there is a point of

non-analyticity in $s(\varepsilon)$, at the energy ε_c of the second order phase transition, this means that the process of maximization of the analytic function $\tilde{s}(\varepsilon, m)$ leads to a non-analytic function $s(\varepsilon)$. This is due to the bimodal shape of $\tilde{s}(\varepsilon, m)$, see Fig. 30. This bimodal shape would not be possible in a short-range system, where $\tilde{s}(\varepsilon, m)$ has to be concave with respect to m at each value ε. This argument has been recently used by Kastner and collaborators [177], with particular emphasis on the ϕ^4 model [166,167] and on the mean-field spherical model [179–182].

In Section 4.1.1, we have emphasized only the concavity with respect to ε. However, in short-range systems the entropy is also concave with respect to other macroscopic variables, e.g. magnetization [183]. Of course, there might be ranges where the concavity is not strict, associated to first order phase transitions. This means that, for short-range systems, a possible non-analyticity in $s(\varepsilon)$ could not derive from the maximization of an analytic $\tilde{s}(\varepsilon, m)$, but that already the function $\tilde{s}(\varepsilon, m)$, over which the maximization is performed, is non-analytic. The mechanism that can give rise to non-analyticities of the entropy in short-range systems has been emphasized by Pettini and coworkers (see [176,178] and references therein). These authors prove [184–186] that a *necessary* condition for having a phase transition in short-range systems is the presence of topological changes in the configurational space subset $\mathcal{M}(v) = (\{q_i\}|U(\{q_i\})/N \leq v)$, which is nothing but the level set v of the potential energy U. This topological change, according to Morse theory, is signaled by the presence of "critical points" in $\mathcal{M}(v)$, i.e. points where the differential of U vanishes. This idea that phase transitions have an origin in some topological and analytical properties of the phase space could of course be extended to systems with long-range interactions. Indeed, it has been extended, and detailed calculations of the critical points of the HMF model [187,188] and of the k-trigonometric model [189–191] have been performed, allowing to show analytically that a topological invariant (the Euler characteristic) has a jump at the phase transition point. However, results in this field raise apparent paradoxes. Indeed, by considering the mean-field ϕ^4 model with varying strength of the mean-field term, Kastner [166,167] prove that nothing special appears in the analytical and topological properties of the potential energy at the second-order phase transition point. Hence, the mechanism here at work to create a phase transition must be different and it is natural to think that it is indeed related to the extremalization of a thermodynamic analytic function (entropy or free energy), independently of the presence of critical points in the potential.

However, although phase transitions in long-range systems, as just explained, can derive also from a mechanism that is not present in short-range systems, still they could also derive from topological changes. The problem to solve is to select which of the "critical points", whose number grows with N, will turn out to determine the few non analytic points of the thermodynamic functions. Recently, it has been shown that the non-analyticities due to "critical points" can become infinitely weak in the thermodynamic limit [182] (solving also a contradiction arisen in a study of the spherical model [192,193]), i.e., the order of the derivative of the entropy which is discontinuous becomes larger and larger when N grows. In order to explain how a discontinuity in a first or second order derivative of the entropy can survive in the thermodynamic limit, Kastner [194] introduces a criterion of "sufficient flatness" of the critical point. An application to the HMF and to the k-trigonometric model confirms this criterion, selecting the correct phase transition point in the thermodynamic limit [194].

All these results concern *necessary* conditions and the much harder problem of finding *sufficient conditions* has not yet been solved. For instance, for a generalization of the k-trigonometric model, it has been found [191] that a phase transition occurs at an energy where no topological change of the phase space manifold is present.

5. Out-of-equilibrium dynamics: Quasi-stationary states

In this section, we will give a brief introduction to the kinetic equations used for long-range systems, i.e. Klimontovich, Vlasov and Lenard–Balescu equations. We will present the numerical evidence of the existence of *quasi-stationary* states in the N-particle dynamics of the Hamiltonian Mean Field (HMF) model. These states, that are out-of-equilibrium and have a lifetime that increases with system size N, turn out to be related to stable stationary states of the Vlasov equation. Selecting the dynamics of a single particle and treating the rest as a bath, we will find that the one-particle velocity distribution function obeys a Fokker–Planck equation with nonlinear diffusion coefficient. Solving this equation, we find cases where diffusion can be normal or otherwise anomalous. A theory, introduced by Lynden-Bell for gravitational systems, which relies on a "fermionic" entropy, allows us finally to predict the one-particle distribution function in the quasi-stationary states for a specific class of initial conditions.

5.1. Kinetic equations

5.1.1. Klimontovich, Vlasov and Lenard–Balescu equations

Here, we will discuss the equations that describe the evolution of the one-particle distribution function for systems with long-range interactions.

The first kinetic equation for a Hamiltonian N-body system was derived by Boltzmann [195]. Boltzmann's theory, in which particles interact only through binary collisions, describes diluted gases with short-range interactions. In order to take into account collisions determined by long-range Coulomb and gravitational forces, Landau [196] and Chandrasekhar [197] modified Boltzmann's collisional term. Mean-field collective effects were first considered by Vlasov [198] and led to the Vlasov equation, later refined by Landau himself [199]. A treatment of collision terms in the context of the Vlasov–Landau approach was developed by Lenard and Balescu [200,201] (see Ref. [202] for a review): collective effects were effectively

taken into account by a dielectric function. These theoretical approaches have also led to the development of kinetic theories for point vortices in two-dimensional hydrodynamics [34,73] and for non-neutral plasmas confined by a magnetic field [203,204]. A detailed recent discussion of the derivation of Landau, Vlasov and Lenard–Balescu equations and of the approximations involved can be found in Ref. [205].

The relation between the N-particle dynamics and the solution of the Vlasov equation has been studied for mean-field systems by Neunzert [206,207] and Braun and Hepp [208,209]. The latter proved that, for sufficiently smooth potentials, the distance between two initially close solutions of the Vlasov equation increases at most exponentially in time. If we apply this result to a large N particle approximation of a continuous distribution the error at $t = 0$ is typically of order $1/\sqrt{N}$, thus for any ϵ and any particle number N, there is a time t up to which the dynamics of the original Hamiltonian and its Vlasov description coincide within an error bounded by ϵ. The theorem implies that this time t increases at least as $\ln N$.

For what concerns the methods to derive the Vlasov equation and higher order kinetic equations, two main perturbative approaches have been followed. The first one begins with the Liouville equation and, either by the usual BBGKY expansion [10,210] or by some projector operator formalism [211,212] leads to the kinetic equation for the one-particle distribution function making some particle decorrelation hypothesis. The second approach is based on the Klimontovich equation [213–215]. In both approaches, the small parameter of the perturbative expansion turns out to be $1/N$, where N is the number of particles. These two methods give completely equivalent results. In the following, we will discuss the method based on the Klimontovich equation. The Vlasov equation will come out at leading order, while the Lenard–Balescu correction term will appear at order $1/N$.

5.1.2. Derivation of Klimontovich equation

The general Hamiltonian we will discuss is of the form

$$H_N = \sum_{j=1}^{N} \frac{P_j^2}{2} + U(\{\Theta_j\}), \tag{259}$$

where we have used the coordinate Θ_j for particle j because, in the following, we will mainly treat models where positions are specified by an angle and

$$U(\Theta_1, \ldots, \Theta_N) = \sum_{i<j}^{N} V(\Theta_i - \Theta_j). \tag{260}$$

We do not use here Kac's scaling of the potential (which amounts here to put a $1/N$ prefactor in the above formula), because we want to maintain the derivation at the general level, which includes both short and long-range interactions. When relevant, we will make comments on where Kac's scaling would produce effects for long-range interactions.

The state of the N-particle system can be described by the *discrete* one-particle time-dependent density function

$$f_d(\theta, p, t) = \frac{1}{N} \sum_{j=1}^{N} \delta\left(\theta - \Theta_j(t)\right) \delta\left(p - P_j(t)\right), \tag{261}$$

where δ is the Dirac function, (θ, p) the Eulerian coordinates of the phase space and (Θ_i, P_i) the Lagrangian coordinates of the N-particles whose dynamics is given by the $2N$ equations of motion

$$\dot{\Theta}_j = P_j, \tag{262}$$

$$\dot{P}_j = -\frac{\partial U}{\partial \Theta_j}. \tag{263}$$

Differentiating with respect to time the one-particle density (261) and using Eqs. (262) and (263), one gets

$$\frac{\partial f_d(\theta, p, t)}{\partial t} = -\frac{1}{N} \sum_j P_j \frac{\partial}{\partial \theta} \delta\left(\theta - \Theta_j(t)\right) \delta\left(p - P_j(t)\right) + \frac{1}{N} \sum_j \frac{\partial U}{\partial \Theta_j} \frac{\partial}{\partial p} \delta\left(\theta - \Theta_j(t)\right) \delta\left(p - P_j(t)\right). \tag{264}$$

Taking advantage of the property of the Dirac δ-function, $a\delta(a - b) = b\delta(a - b)$, it is possible to rewrite this equation as

$$\frac{\partial f_d(\theta, p, t)}{\partial t} = -\frac{1}{N} \sum_j p \frac{\partial}{\partial \theta} \delta\left(\theta - \Theta_j(t)\right) \delta\left(p - P_j(t)\right) + \frac{1}{N} \sum_j \frac{\partial v}{\partial \theta} \frac{\partial}{\partial p} \delta\left(\theta - \Theta_j(t)\right) \delta\left(p - P_j(t)\right), \tag{265}$$

where

$$v(\theta, t) = N \int \mathrm{d}\theta' \mathrm{d}p' V(\theta - \theta') f_d(\theta', p', t), \tag{266}$$

which leads to the Klimontovich equation

$$\frac{\partial f_d}{\partial t} + p\frac{\partial f_d}{\partial \theta} - \frac{\partial v}{\partial \theta}\frac{\partial f_d}{\partial p} = 0. \tag{267}$$

One notes that the above derivation is *exact* even for a finite number of particles N. However, this equation contains the information about the orbit of every single particle (since f_d depends on the $2N$ Lagrangian coordinates of each particle, (Θ_i, P_i)) which is far more than we need. Hence, the Klimontovich equation is especially useful as a starting point for the derivation of approximate equations that describe the average properties of the system. This is what we discuss in the following subsections.

The more widely used approach to kinetic equations begins with the Liouville equation. This equation governs the time evolution of the probability density in the full $2N$ dimensional phase space $(\theta_1, p_1, \ldots, \theta_N, p_N)$. To arrive at the Liouville equation, let us first define the Klimontovich distribution function

$$F_d\left(\theta_1, p_1, \theta_2, p_2, \ldots, \theta_N, p_N, t\right) = \prod_{j=1}^{N} \delta\left(\theta_i - \Theta_j\left(t\right)\right) \delta\left(p_j - P_j\left(t\right)\right), \tag{268}$$

where $(\Theta_i(t), P_i(t))$ are again the Lagrangian coordinates. Deriving (268) with respect to time, and using the equations of motion, one obtains

$$\frac{\partial F_d}{\partial t} + \sum_{i=1}^{N} p_i \frac{\partial F_d}{\partial \theta_i} - \sum_{i=1}^{N} \frac{\partial U(\theta_1, \ldots, \theta_N)}{\partial \theta_i}\frac{\partial F_d}{\partial p_i} = 0. \tag{269}$$

This is the Liouville equation for F_d. It has been derived for a particular distribution function, which describes a single point in the $2N$ dimensional space. The Liouville equation for a generic distribution function is obtained by an averaging procedure which leads to the introduction of a smooth density $\rho(\theta_1, p_1, \theta_2, p_2, \ldots, \theta_N, p_N, t)$, which can be shown to obey the same Eq. (269). This smoothing procedure is similar to the one that leads to the Vlasov equation, that we will describe just below. However, while the Liouville equation for ρ is exact (as the Klimontovich equation for f_d), the Vlasov equation is not exact and we will indeed discuss finite N corrections. It is well known that the fluid described by the density ρ is incompressible.

5.1.3. Vlasov equation: Collisionless approximation of the Klimontovich equation

Determining $f_d(\theta, p, t)$, which characterizes whether a point particle is to be found at a given point (θ, p) in Eulerian phase space, would imply solving the equations of motion (262) and (263) with initial conditions $(\{\Theta_i(0), P_i(0)\})$. In general, this is a difficult task and typically not feasible for nonlinear systems. Alternatively, one can define an averaged one-particle density function using an infinite number of realizations prepared according to some prescription. One could for instance consider a large number of initial conditions, close to the same macroscopic state. Let us define the density of such initial macroscopic state as $f_{\rm in}(\{\Theta_i(0), P_i(0)\})$. The average one-particle density function f_0 is obtained as

$$f_0(\theta, p, t) \equiv \langle f_d(\theta, p, t)\rangle = \int \prod_i \mathrm{d}\Theta_i(0)\mathrm{d}P_i(0) f_{\rm in}(\{\Theta_i(0), P_i(0)\}) f_d(\theta, p, t), \tag{270}$$

where the dependence of f_d on $(\{\Theta_i(0), P_i(0)\})$ comes from the solution of the equations of motion that enter the definition of f_d, see Eq. (261).

The equation for the time evolution of the smoothed distribution f_0 is obtained by again averaging over $f_{\rm in}$. Before doing it, let us introduce the definition of the fluctuations δf around the smooth distribution

$$f_d(\theta, p, t) = f_0(\theta, p, t) + \frac{1}{\sqrt{N}}\delta f(\theta, p, t). \tag{271}$$

Evidently f_0 is independent of the detailed microscopic properties of the initial state and depends on it only through $f_{\rm in}$, while δf, like f_d depends on all Lagrangian variables of the initial state. Obviously, the average over $f_{\rm in}$ of δf is zero. The introduction of the prefactor $1/\sqrt{N}$ takes into account the typical size of relative fluctuations. Indeed, $f_d - f_0$ is the difference between a singular distribution, containing Dirac deltas, and a smooth function; therefore the statement that this difference is of the order $1/\sqrt{N}$ has to be interpreted physically. Its meaning is the following: if we integrate $f_d - f_0$ in θ and p in a volume which is small compared to the total available volume, but large enough to contain many particles, then the value of this integral, both at equilibrium and out-of-equilibrium, is of the order $1/\sqrt{N}$.

Inserting Eq. (271) into Eq. (266) leads to

$$v(\theta, t) = \langle v\rangle(\theta, t) + \frac{1}{\sqrt{N}}\delta v(\theta, t), \tag{272}$$

where the first term comes from the average over $f_{\rm in}$, and coincides with

$$\langle v\rangle(\theta, t) = N\int \mathrm{d}\theta'\mathrm{d}p' V(\theta - \theta') f_0(\theta', p', t), \tag{273}$$

while the second term defines δv, which depends on all the details of the initial state. If we had used Kac's scaling the N prefactor in formula (273) would be absent, making the average potential intensive. This latter scaling is commonly used [209] and would be more appropriate to derive the Boltzmann equation.

Inserting both expressions (271) and (272) in the Klimontovich equation (267), one obtains

$$\frac{\partial f_0}{\partial t}+p\frac{\partial f_0}{\partial \theta}-\frac{\partial \langle v\rangle}{\partial \theta}\frac{\partial f_0}{\partial p}=-\frac{1}{\sqrt{N}}\left(\frac{\partial \delta f}{\partial t}+p\frac{\partial \delta f}{\partial \theta}-\frac{\partial \delta v}{\partial \theta}\frac{\partial f_0}{\partial p}-\frac{\partial \langle v\rangle}{\partial \theta}\frac{\partial \delta f}{\partial p}\right)+\frac{1}{N}\frac{\partial \delta v}{\partial \theta}\frac{\partial \delta f}{\partial p}. \tag{274}$$

Taking the average over $f_{\rm in}$ of this equation leads to

$$\frac{\partial f_0}{\partial t}+p\frac{\partial f_0}{\partial \theta}-\frac{\partial \langle v\rangle}{\partial \theta}\frac{\partial f_0}{\partial p}=\frac{1}{N}\left\langle\frac{\partial \delta v}{\partial \theta}\frac{\partial \delta f}{\partial p}\right\rangle. \tag{275}$$

It is important to stress that above equation is still exact. Up to now we have not even made any hypothesis about the long or short-range properties of the potential V. In standard kinetic theory, Eq. (275) would correspond to the first equation of the BBGKY hierarchy (see e.g. Chapter 4 of Ref. [215]). For short-range interactions, the r.h.s. of this equation would originate the leading contribution to the collision term of the Boltzmann equation, while the third term of the l.h.s. would be negligible close to equilibrium. On the contrary, for long-range interactions, it will turn out that the r.h.s is of order $1/N$ [196,200,201]. Indeed, the scaling with $1/\sqrt{N}$ in Eqs. (271) and (272) is appropriate for long-range interactions. Therefore, for long-range interactions, the third term on l.h.s. will be the leading term and, in the limit $N\to\infty$, one ends up with the *Vlasov equation*

$$\frac{\partial f_0}{\partial t}+p\frac{\partial f_0}{\partial \theta}-\frac{\partial \langle v\rangle}{\partial \theta}\frac{\partial f_0}{\partial p}=0. \tag{276}$$

To distinguish the effects due to discreteness (finite value of N) from the *collective* effects grouped on the l.h.s. of Eq. (275), the r.h.s. is usually referred to as *collisional* term. This is the origin of the name *collisionless Boltzmann equation* sometimes used for the Vlasov equation. It is however important to underline that there are no true collisions for long-range systems: granular effects, discreteness effects or finite N corrections would be more appropriate names.

The Vlasov equation has wide applications in gravitational systems [22] and in plasma physics [215]. A typical question that is posed is the one of stability of the stationary solutions. In particular, the stability of stationary homogeneous solutions has important applications in plasma physics for the description of Landau damping of for beam-plasma instabilities. In the following subsection, we will discuss in some detail the stability of stationary homogenous solution for the Hamiltonian Mean Field model.

5.1.4. Stationary stable solutions of the Vlasov equation: Application to the HMF model

Let us concentrate on homogenous states, for which the one-particle distribution function does not depend on θ, so that $f_0=f_0(p,t)$. For these distributions, using Eq. (273), one immediately gets that $\langle v\rangle=const.$, and consequently the Vlasov equation (276) reduces to

$$\frac{\partial f_0}{\partial t}(p,t)=0. \tag{277}$$

Therefore, homogeneous distributions are stationary. However, stationarity does not imply *stability*. Indeed, the stability of stationary spatially homogeneous solutions can be studied by subtracting Eq. (275) from Eq. (274). One gets

$$\frac{\partial \delta f}{\partial t}+p\frac{\partial \delta f}{\partial \theta}-\frac{\partial \delta v}{\partial \theta}\frac{\partial f_0}{\partial p}-\frac{\partial \langle v\rangle}{\partial \theta}\frac{\partial \delta f}{\partial p}=\frac{1}{\sqrt{N}}\left[\frac{\partial \delta v}{\partial \theta}\frac{\partial \delta f}{\partial p}-\left\langle\frac{\partial \delta v}{\partial \theta}\frac{\partial \delta f}{\partial p}\right\rangle\right]. \tag{278}$$

For times much shorter than $\sqrt{N}$ (or equivalently for $N\to\infty$), we may drop the r.h.s. of Eq. (278), which contains quadratic terms in the fluctuations. As the last term of the l.h.s. vanishes, since $\langle v\rangle=const.$, the fluctuating part of f_d, δf, obeys the *linearized Vlasov equation* [198,199,208,209,216]

$$\frac{\partial \delta f}{\partial t}+p\frac{\partial \delta f}{\partial \theta}-\frac{\partial \delta v}{\partial \theta}\frac{\partial f_0}{\partial p}=0. \tag{279}$$

This equation could also be easily obtained by linearizing directly the Vlasov equation (276). Looking for plane wave solutions

$$\delta f(\theta,p,t)=\hat{f}(p)e^{i(k\theta-\omega t)} \tag{280}$$

$$\delta v(\theta,t)=\hat{A}e^{i(k\theta-\omega t)}, \tag{281}$$

we obtain

$$-i\omega\hat{f}(p)+pik\hat{f}(p)-ik\hat{A}f_0'(p)=0 \tag{282}$$

which leads to

$$\hat{f}(p) = \frac{k\hat{A}}{pk-\omega} f_0'(p). \tag{283}$$

This analysis is completely general and applies to all one-dimensional models.

To be more specific, let us consider the Hamiltonian Mean Field model (126). In this particular case, the potential appearing in the Vlasov equation (276) is

$$\langle v\rangle(\theta,t) = \int_0^{2\pi} \mathrm{d}\alpha \int_{-\infty}^{+\infty} \mathrm{d}p[1-\cos(\theta-\alpha)]f_0(\alpha,p,t). \tag{284}$$

We now use Eqs. (280), (281) and (283), where for the HMF model k is an integer because the potential is 2π-periodic, in the definition of the fluctuations of the potential

$$\delta v(\theta,t) = \int_0^{2\pi} \mathrm{d}\alpha \int_{-\infty}^{+\infty} \mathrm{d}p[1-\cos(\theta-\alpha)]\delta f(\alpha,p,t). \tag{285}$$

We obtain

$$\delta v(\theta,t) = \hat{A}\mathrm{e}^{\mathrm{i}(k\theta-\omega t)} = -\pi k\left(\delta_{k,1}+\delta_{k,-1}\right)\int_{-\infty}^{+\infty} \mathrm{d}p \frac{f_0'(p)}{pk-\omega}\hat{A}\mathrm{e}^{\mathrm{i}(k\theta-\omega t)}. \tag{286}$$

Introducing the so-called "plasma response dielectric function" (see Chapter 6 of Ref. [215])

$$\tilde{D}(\omega,k) \equiv 1+\pi k\left(\delta_{k,1}+\delta_{k,-1}\right)\int_{-\infty}^{+\infty} \mathrm{d}p \frac{f_0'(p)}{pk-\omega}, \tag{287}$$

Eq. (286) can be rewritten as $\tilde{D}(\omega,k)\hat{A}\exp(\mathrm{i}(k\theta-\omega t)) = 0$. In order to get non-vanishing solutions for $\hat{A}$, the relation

$$\tilde{D}(\omega,k) = 0 \tag{288}$$

must be satisfied, which leads to the dispersion relation $\omega = \omega(k)$, linking the frequency with the wavevector. From the definition (287), it is evident that the only possible collective modes are $k = \pm 1$. The denominator in the integral on the r.h.s of Eq. (287) must be treated with care. Indeed, a difficulty arises if one attempts to obtain solutions $\omega(k)$ of Eq. (288) corresponding to purely oscillatory solutions of the linearized Vlasov equation (279), i.e. with $\mathrm{Im}(\omega) = 0$. These solutions are important because they lie at the boundary between stable and unstable modes, which correspond to negative and positive values of $\mathrm{Im}(\omega)$, respectively. If one makes the substitution of real values of ω into the integral mentioned above, one notes that the integrand has a pole at $p = \omega/k$ and therefore the integral is not well defined. The prescription for performing this singular integral is to deform the integration contour of p in such a way to circulate around the pole $p = \omega/k$. This method was introduced in plasma physics by Landau [199] and the deformed contour is called *Landau contour*. The deformation of the contour is equivalent to the displacement of the pole $\omega \to \omega \pm \mathrm{i}0$. More rigorously this is expressed by the Plemelj formula

$$\lim_{\gamma\to 0}\frac{1}{x-a\pm\mathrm{i}|\gamma|} = \mathcal{P}\frac{1}{x-a} \mp \mathrm{i}\pi\delta(x-a) \tag{289}$$

where $\mathcal{P}$ denotes the principal value, which applies when integrating over x. The second term on the r.h.s comes from integrating on a semi-circle around the pole. The direction to be chosen for the displacement of the pole is determined if one uses the Laplace transform in time, suited for the solution of the initial value problem associated to our kinetic equation. We will use this technique in the next subsection, treating the Lenard–Balescu equation. The Laplace transform is defined for $\mathrm{Im}(\omega) > 0$. From this, we derive that $\tilde{D}(\omega,k)$ is defined by Eq. (287) for $\mathrm{Im}(\omega) > 0$, and by its analytic continuation for $\mathrm{Im}(\omega) \le 0$. In particular, this means that for $\mathrm{Im}(\omega) = 0$ the rule for encircling the singularity at $p = \omega/k$ is obtained by the Plemelj formula (289) by posing $\omega \to \omega + \mathrm{i}0$. Applying this rule, we find that for $k = \pm 1$ the dielectric function for real values of ω is given by

$$\tilde{D}(\omega,\pm 1) = 1+\pi\mathcal{P}\int_{-\infty}^{+\infty} \mathrm{d}p \frac{f_0'(p)}{p\mp\omega} \pm \mathrm{i}\pi^2 f_0'(\pm\omega). \tag{290}$$

We see in particular that $\tilde{D}(-\omega,-1) = \tilde{D}^*(\omega,1)$. As we mentioned, the real solutions $\omega(k)$ obtained from equating to zero the dielectric function correspond to the boundary between stable and unstable modes. From the last equation, posing equal to zero both the real and the imaginary parts, we see that in order to have a real ω as a solution of $\tilde{D}(\omega,\pm 1) = 0$, it is necessary that $f_0'(\pm\omega) = 0$. If one considers initial distributions with a single maximum at the origin in momentum space, one obtains, for both $\tilde{D}(\omega,1)$ and $\tilde{D}(\omega,-1)$, $\omega = 0$. Hence, the real part of Eq. (290) will be zero if

$$I \equiv 1+\pi\int_{-\infty}^{+\infty} \frac{f_0'(p)}{p}\mathrm{d}p = 0. \tag{291}$$

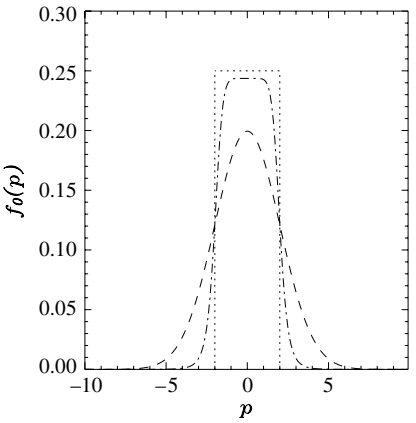


Fig. 31. Three examples of stationary homogenous solutions of the Vlasov equation. The Gaussian (dashed), the water-bag (dotted) and the power-law (Eq. (295)) in the case $\nu = 8$ (dash–dotted).

This gives the stability boundary of the homogeneous distribution $f_0(p)$. To deduce where the stability region is located in the functional space of $f_0(p)$, one can perform an analysis based on the Nyquist criterion, from which one finds that the stability condition is $I > 0$ [135].

The linear stability analysis discussed above appeared in a series of different papers [130,135,217–220], The stability boundary can be also derived using the so-called energy-Casimir method [26].

Condition (291) is a functional equation in $f_0(p)$. By a convenient parametrization of $f_0(p)$, one can obtain a stability boundary in a finite dimensional control parameter space. A relevant parametrization for the HMF model is the one in terms of the energy per particle ε. Therefore, we will briefly discuss below some examples where a critical energy density ε^c appears, above which the homogeneous solution is stable.

- The first one is the Gaussian distribution

$$f_{\rm g}(p) = \frac{1}{2\pi}\sqrt{\frac{\beta}{2\pi}}\,\exp(-\beta p^2/2) \tag{292}$$

(see Fig. 31). One gets $I = 1 - \beta/2$. The condition $I = 0$ coincides with the equilibrium statistical mechanics result that the critical inverse temperature is $\beta_{\rm g}^c = 2$, and its associated critical energy is $\varepsilon_{\rm g}^c = 3/4$ [128,134,135].
- The second example is the water-bag distribution

$$f_{\rm wb}(p) = \frac{1}{2\pi}\frac{1}{2p_0}\left[\Theta(p+p_0) - \Theta(p-p_0)\right] \tag{293}$$

where Θ is the Heaviside function. This distribution, also shown in Fig. 31, has been often used in the past to test numerically the out-of-equilibrium properties of the HMF model. In this case, one obtains $I = 1 - 1/(2p_0^2)$ which leads to a smaller critical energy $\varepsilon_{\rm wb}^c = 7/12$.
- Another example studied in the literature is the q-Gaussian

$$f_{\rm T}(p) \sim [1 - \alpha(1-q)p^2]^{\frac{1}{1-q}}, \tag{294}$$

with α positive. This function has a compact support for $q < 1$, while it decays with a power law tail for $q > 1$; we recover the Gaussian for $q = 1$. The water-bag distribution is recovered for $q \to -\infty$. For $q \geq 5/3$, the second moment of this distribution, and thus the average kinetic energy, is not finite, and therefore a cut-off to keep the energy finite is used in actual implementations. For the cases with finite kinetic energy, i.e., for $q < 5/3$, one gets [130,221,222] $\varepsilon_{\rm T}^c = \frac{3}{4} + \frac{q-1}{2(5-3q)}$. One easily verifies that the previous values of the critical energy for the Gaussian and the water-bag distribution are recovered for the proper values of q.
- The last example [223] is a distribution with power-law tails

$$f_{\rm pl}(p) = \frac{A_\nu}{1 + |p/p_0|^\nu}, \tag{295}$$

characterized by the exponent ν. The unity, added in the denominator to avoid the divergence at the origin $p = 0$, does not affect neither the asymptotic form, nor the theoretical predictions. The parameter $p_0 = \sqrt{\frac{\sin(3\pi/\nu)}{\sin(\pi/\nu)}\frac{K}{N}}$ controls the kinetic energy density K/N and the normalization factor is $A_\nu = \nu\sin(\pi/\nu)/(2\pi p_0)$. The exponent ν must be greater than 3 to get a finite kinetic energy: we have used $\nu = 8$ (see Fig. 31). Note that the power-law distribution cannot be included in the q-exponentials family [224], although it has similar power law tails at large $|p|$. Distribution (295) is

stable above the critical energy density $\varepsilon_{\rm pl}^c = \frac{1}{2} + \frac{\sin(\pi/\nu)}{4\sin(3\pi/\nu)}$, which corresponds to $\varepsilon_{\rm pl}^c = 0.75,\ 0.625$ and $0.60355\ldots$ for $\nu = 4,\ 6$ and 8 respectively.

- Finally let us consider a more general distribution, namely a mixed distribution between $f_{\rm wb}$ and $f_{\rm g}$, defined as

$$f_a(p) = (1-a)f_{\rm wb}(p) + af_{\rm g}(p). \tag{296}$$

Thanks to the linearity of the quantity (291) with respect to the distribution, the critical energy density for this mixed distribution f_a is obtained [26] as a linear combination of the previous results.

The condition (291) defines therefore a critical energy ε^c which is, in general, different from the critical energy 3/4 where the second-order phase transition of equilibrium statistical mechanics is located. However, as expected, the two values coincide for a Gaussian distribution.

All the above distributions are thus stationary solutions of the Vlasov equation (276). However, it is important to realize that they are Vlasov stable stationary solutions among *infinitely* many others and there is no reason to emphasize one more than the others.

5.1.5. The Lenard–Balescu equation

We have so far concentrated on the collisionless dynamics as described by the Vlasov equation. When using this equation, one implicitly assumes that the particles of the system move under the influence of the average potential generated by all the other particles. This means that the acceleration of all the particles is given by a single function, i.e., by $-\partial\langle v\rangle/\partial\theta$. However, this assumption is not valid for arbitrarily long times. The presence of individual "collisions" invalidate the assumption that the acceleration of all the particles derives from a single mean field function: "collisions" perturb particles away from the trajectories they would have taken if the distribution of particles in the system were perfectly smooth. In long-range systems, however, we have seen that for times shorter than N (thus for times that can be very long), the collisionless approximation is valid. For larger times, the Vlasov equation is not valid and we have to consider the effect of "collisions". Under this effect, particles will deviate from the orbits determined by the Vlasov equation on a characteristic time that is called *relaxation time*. This can be justified from the fact that a stable stationary solution of the Vlasov equation, as determined from the analysis described in the previous subsection, will eventually be perturbed by the effect of "collisions".

Let us then go back to the exact equation (275) and consider the right-hand side. At the next level of approximation, i.e., the level $1/N$, we can determine the right-hand side using the solutions for δv and δf as determined by the collisionless dynamics, given by Eq. (279). If again we restrict to the case in which f_0, a stable stationary solution of the Vlasov equation, does not depend on θ, at this level of approximation our kinetic equation will therefore be

$$\frac{\partial f_0}{\partial t} = \frac{1}{N}\left\langle \frac{\partial \delta v}{\partial \theta}\frac{\partial \delta f}{\partial p}\right\rangle. \tag{297}$$

This equation is valid up to a time where the right-hand side makes f_0 leave its stability basin as determined by the Vlasov equation. As already anticipated, one can expect this time to be of order N. This equation is called the Lenard–Balescu equation, and its solution requires the determination of the correlation function on the right-hand side of Eq. (297). To achieve this goal, it is necessary to use the spatio-temporal Fourier–Laplace transform. Using this technique, one is able to determine the correlation expressing the collisional term as a function of its value at the initial time. We thus have to solve an initial value problem.

The spatio-temporal Fourier–Laplace transform of the fluctuation of the density δf is defined by

$$\widetilde{\delta f}(k,p,\omega) = \int_0^{2\pi}\frac{{\rm d}\theta}{2\pi}\int_0^{+\infty}{\rm d}t{\rm e}^{-{\rm i}(k\theta-\omega t)}\delta f(\theta,p,t), \tag{298}$$

associated with a similar expression for the fluctuation of the potential δv. As usual with Laplace transforms, $\widetilde{\delta f}(k,p,\omega)$ is defined by the last equation only for ${\rm Im}(\omega)$ sufficiently large. For the remaining part of the complex ω plane, it is defined by an analytic continuation. The inverse transform is

$$\delta f(\theta,p,t) = \sum_{k=-\infty}^{+\infty}\int_{\mathcal{C}}\frac{{\rm d}\omega}{2\pi}{\rm e}^{{\rm i}(k\theta-\omega t)}\ \widetilde{\delta f}(k,p,\omega), \tag{299}$$

where the Laplace contour $\mathcal{C}$ in the complex ω plane must pass above all poles of the integrand. The inverse transform has a sum over the integer values of k since the coordinates θ take values from 0 to 2π. In fact, we will determine the right hand side of Eq. (297) taking as an example the HMF model, as we have done in the previous subsection. More generally, there will be an integral over k, but the analysis will be perfectly equivalent.

If we multiply Eq. (279) by ${\rm e}^{-{\rm i}(k\theta-\omega t)}$ and integrate over θ from 0 to 2π and over t from 0 to ∞, we obtain

$$-\widehat{\delta f}(k,p,0) - {\rm i}\omega\,\widetilde{\delta f}(k,p,\omega) + {\rm i}kp\,\widetilde{\delta f}(k,p,\omega) - {\rm i}k\,\widetilde{\delta v}(k,\omega)f_0'(p) = 0, \tag{300}$$

where the first term is the spatial Fourier transform of the initial value

$$\widehat{\delta f}(k,p,0) = \int_0^{2\pi}\frac{{\rm d}\theta}{2\pi}{\rm e}^{-{\rm i}k\theta}\delta f(\theta,p,0), \tag{301}$$

and it arises from the integration by parts in obtaining the Laplace–Fourier transform of $\partial\delta f/\partial t$. The above equation can be rewritten as

$$\widetilde{\delta f}(k,p,\omega)=\frac{kf_0'(p)}{pk-\omega}\,\widetilde{\delta v}(k,\omega)+\frac{\widehat{\delta f}(k,p,0)}{\mathrm{i}(pk-\omega)}, \tag{302}$$

where one identifies a first "collective" term depending on the perturbation of the potential, and a second one which depends on the initial condition. Combining Eq. (302) with the Laplace–Fourier transform of the fluctuations of the potential (285), i.e.

$$\widetilde{\delta v}(k,\omega)=-\pi\left(\delta_{k,1}+\delta_{k,-1}-2\delta_{k,0}\right)\int_{-\infty}^{+\infty}\mathrm{d}p\;\widetilde{\delta f}(k,p,\omega), \tag{303}$$

and integrating (302) over the p variable, gives

$$\int_{-\infty}^{+\infty}\mathrm{d}p\;\widetilde{\delta f}(k,p,\omega)\left[1+\pi k\left(\delta_{k,1}+\delta_{k,-1}\right)\int_{-\infty}^{+\infty}\mathrm{d}p'\frac{f_0'(p')}{(p'k-\omega)}\right]=\int_{-\infty}^{+\infty}\mathrm{d}p\frac{\widehat{\delta f}(k,p,0)}{\mathrm{i}(pk-\omega)}, \tag{304}$$

where one recognizes the plasma response dielectric function (287) in the parenthesis. Eq. (303) can thus be rewritten as

$$\widetilde{\delta v}(k,\omega)=-\frac{\pi\left(\delta_{k,1}+\delta_{k,-1}-2\delta_{k,0}\right)}{\tilde{D}(\omega,k)}\int_{-\infty}^{+\infty}\mathrm{d}p\frac{\widehat{\delta f}(k,p,0)}{\mathrm{i}(pk-\omega)}. \tag{305}$$

We see that the Laplace contour $\mathcal{C}$ for the inversion formula must pass above all zeroes of $\tilde{D}(k,\omega)$. We can consider that these zeros will all be located in the half-plane $\mathrm{Im}(\omega)<0$, since otherwise the problem of the $1/N$ perturbations to a stable stationary solution of the Vlasov equation would not make sense [214].

One can then use these expressions to compute the collisional term appearing on the right-hand-side of Eq. (297). Forgetting temporarily the factor $1/N$ and the derivative with respect to the variable p, one has

$$\left\langle\frac{\partial\delta v}{\partial\theta}\delta f\right\rangle=\left\langle\sum_{k=-\infty}^{+\infty}\int_{\mathcal{C}}\frac{\mathrm{d}\omega}{2\pi}\mathrm{i}k\mathrm{e}^{\mathrm{i}(k\theta-\omega t)}\widetilde{\delta v}(k,\omega)\sum_{k'=-\infty}^{+\infty}\int_{\mathcal{C}'}\frac{\mathrm{d}\omega'}{2\pi}\mathrm{e}^{\mathrm{i}(k'\theta-\omega' t)}\;\widetilde{\delta f}(k',p,\omega')\right\rangle \tag{306}$$

$$=\frac{1}{(2\pi)^2}\sum_{k=-\infty}^{+\infty}\sum_{k'=-\infty}^{+\infty}\int_{\mathcal{C}}\mathrm{d}\omega\int_{\mathcal{C}'}\mathrm{d}\omega'\mathrm{i}k\mathrm{e}^{\mathrm{i}[(k+k')\theta-(\omega+\omega')t]}\left\langle\widetilde{\delta v}(k,\omega)\,\widetilde{\delta f}(k',p,\omega')\right\rangle, \tag{307}$$

which relies on evaluating the correlation $\langle\widetilde{\delta v}(k,\omega)\,\widetilde{\delta f}(k',p,\omega')\rangle$. The presence of the factor k in the last equation allows us to forget the $k=0$ contribution to $\widetilde{\delta v}(k,\omega)$ in Eq. (305); this was expected, since the constant component of δv cannot contribute to the force. Using Eq. (302), one finds

$$\left\langle\widetilde{\delta v}(k,\omega)\,\widetilde{\delta f}(k',p,\omega')\right\rangle=\frac{k'f_0'(p)}{pk'-\omega'}\left\langle\widetilde{\delta v}(k,\omega)\,\widetilde{\delta v}(k',\omega')\right\rangle+\frac{\left\langle\widetilde{\delta v}(k,\omega)\,\widehat{\delta f}(k',p,0)\right\rangle}{\mathrm{i}(pk'-\omega')}. \tag{308}$$

The first term on the r.h.s. corresponds to the self-correlation of the potential, while the second one to the correlation between the fluctuations of the potential and of the distribution at time $t=0$. Let us consider separately the two terms of the last equation.

From Eq. (305), neglecting the $k=0$ contribution and taking the statistical average, we have

$$\langle\widetilde{\delta v}(k,\omega)\widetilde{\delta v}(k',\omega')\rangle=\frac{\pi^2\left(\delta_{k,1}+\delta_{k,-1}\right)\left(\delta_{k',1}+\delta_{k',-1}\right)}{\tilde{D}(\omega,k)\tilde{D}(\omega',k')}\int_{-\infty}^{+\infty}\mathrm{d}p\int_{-\infty}^{+\infty}\mathrm{d}p'\frac{\langle\widehat{\delta f}(k,p,0)\widehat{\delta f}(k',p',0)\rangle}{\mathrm{i}(pk-\omega)\mathrm{i}(p'k'-\omega')} \tag{309}$$

$$=\frac{\pi}{2}\frac{\delta_{k,-k'}\left(\delta_{k,1}+\delta_{k,-1}\right)}{\tilde{D}(\omega,k)\tilde{D}(\omega',-k)}\int_{-\infty}^{+\infty}\mathrm{d}p\int_{-\infty}^{+\infty}\mathrm{d}p'\frac{\left[f_0(p)\delta(p-p')+\mu(k,p,p')\right]}{(pk-\omega)(p'k+\omega')}, \tag{310}$$

where we have replaced the autocorrelation of the fluctuation of the distribution at $t=0$ (see Appendix D) by

$$\langle\widehat{\delta f}(k,p,0)\widehat{\delta f}(k',p',0)\rangle=\frac{\delta_{k,-k'}}{2\pi}\left[f_0(p)\delta(p-p')+\mu(k,p,p')\right], \tag{311}$$

with the first term expressing the single particle contribution to the correlation, while in the second term the function $\mu(k,p,p')$ comes from the contribution of different particles. This last function is smooth, but otherwise arbitrary, since it is related to initial conditions of our initial value problem. However, it can be shown [214] that, going back in the time domain with the inverse Laplace–Fourier transform, the contribution of this function to the correlation decays in time. Therefore we can consider only the first term in Eq. (311), obtaining

$$\langle\widetilde{\delta v}(k,\omega)\widetilde{\delta v}(k',\omega')\rangle=\frac{\pi}{2}\frac{\delta_{k,-k'}\left(\delta_{k,1}+\delta_{k,-1}\right)}{\tilde{D}(\omega,k)\tilde{D}(\omega',-k)}\int_{-\infty}^{+\infty}\mathrm{d}p\;\frac{f_0(p)}{(pk-\omega)(pk+\omega')}. \tag{312}$$

Considering again only the contributions that, after integration in ω and ω', do not decay in time, it can be shown [214], through the use of the Plemelj formula (289), that $[(pk-\omega)(pk+\omega')]^{-1}$ can be substituted by $(2\pi)^2\delta(\omega+\omega')\delta(\omega-pk)$. In addition, using the property $\tilde{D}(\omega,k)=\tilde{D}^*(-\omega,-k)$, one finally ends up with the result

$$\left\langle \delta v(k,\omega)\delta v(k',\omega')\right\rangle = 2\pi^3\delta_{k,-k'}\left(\delta_{k,1}+\delta_{k,-1}\right)\frac{\delta(\omega+\omega')}{\left|\tilde{D}(\omega,k)\right|^2}\int \mathrm{d}p f_0(p)\delta(\omega-pk), \tag{313}$$

which vanishes except for $|k|=1$: only the cases $k=-k'=\pm 1$ therefore contribute.

We now consider the second term of Eq. (308). Using again Eq. (305) without the $k=0$ term we have

$$\frac{\left\langle \widetilde{\delta v}(k,\omega)\,\widehat{\delta f}(k',p,0)\right\rangle}{\mathrm{i}(pk'-\omega')} = -\frac{\pi\left(\delta_{k,1}+\delta_{k,-1}\right)}{\tilde{D}(\omega,k)}\int_{-\infty}^{+\infty}\mathrm{d}p'\,\frac{\left\langle \widehat{\delta f}(k,p',0)\,\widehat{\delta f}(k',p,0)\right\rangle}{\mathrm{i}(p'k-\omega)\mathrm{i}(pk'-\omega')}. \tag{314}$$

As for the analysis of the first term of Eq. (308), we substitute the initial time correlation in the last integral with the first term in Eq. (311), and afterwards we substitute $[(pk-\omega)(pk+\omega')]^{-1}$ with $(2\pi)^2\delta(\omega+\omega')\delta(\omega-pk)$. We therefore obtain

$$\frac{\left\langle \widetilde{\delta v}(k,\omega,k)\,\widehat{\delta f}(k',p,0)\right\rangle}{\mathrm{i}(k'p-\omega')} = -\frac{2\pi^2\delta_{k,-k'}\left(\delta_{k,1}+\delta_{k,-1}\right)}{\tilde{D}(\omega,k)}f_0(p)\delta(\omega+\omega')\delta(\omega-pk). \tag{315}$$

From Eq. (313), one gets the contribution to (307) of the first term of Eq. (308). Exploiting the presence of the factors $\delta_{k,-k'}$ and $\delta(\omega+\omega')$, this contribution is

$$\frac{\mathrm{i}\pi}{2}\sum_{k=-\infty}^{+\infty}\int_{\mathcal{C}}\mathrm{d}\omega\left(\delta_{k,1}+\delta_{k,-1}\right)\frac{k^2f_0'(p)}{pk-\omega}\frac{1}{\left|\tilde{D}(\omega,k)\right|^2}\int\mathrm{d}p' f_0(p')\delta(\omega-p'k)$$

$$= \frac{\mathrm{i}\pi}{2}\int_{\mathcal{C}}\mathrm{d}\omega f_0'(p)\left[\frac{1}{p-\omega}\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}-\frac{1}{p+\omega}\frac{f_0(-\omega)}{\left|\tilde{D}(\omega,-1)\right|^2}\right] \tag{316}$$

$$= \frac{\mathrm{i}\pi}{2}\int_{-\infty}^{+\infty}\mathrm{d}\omega f_0'(p)\left[\left(\mathcal{P}\frac{1}{p-\omega}-\mathrm{i}\pi\delta(p-\omega)\right)\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}-\left(\mathcal{P}\frac{1}{p+\omega}+\mathrm{i}\pi\delta(p+\omega)\right)\frac{f_0(-\omega)}{\left|\tilde{D}(\omega,-1)\right|^2}\right] \tag{317}$$

$$= \pi^2\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{1}{\left|\tilde{D}(\omega,1)\right|^2}f_0'(p)f_0(\omega)\delta(p-\omega). \tag{318}$$

Passing from the second to the third line we have used again the Plemelj formula, that allows us to integrate on the real ω axis. Note that in this case the rule of singularity encircling is the opposite of the usual one, i.e., it is $\omega\to\omega-\mathrm{i}0$, since the ω of the $(pk-\omega)$ term in the denominator in the first line comes from the integration in ω', that gives $\omega=-\omega'$ [214]. Passing from the third to the fourth line, the variable of integration has been changed from ω to $-\omega$ in the last two terms of the third line, and we have also used the property $\tilde{D}(\omega,k)=\tilde{D}^*(-\omega,-k)$. The contribution of the second term of (308) to Eq. (307) is obtained from (315). Exploiting again the factors $\delta_{k,-k'}$ and $\delta(\omega+\omega')$, this contribution is

$$\frac{-\mathrm{i}}{2}\sum_{k=-\infty}^{+\infty}\int_{\mathcal{C}}\mathrm{d}\omega f_0(p)\left(\delta_{k,1}+\delta_{k,-1}\right)\frac{\delta(\omega-pk)}{\tilde{D}(\omega,k)} = \frac{-\mathrm{i}}{2}\int_{-\infty}^{+\infty}\mathrm{d}\omega f_0(p)\frac{\delta(\omega-p)}{\left|\tilde{D}(\omega,1)\right|^2}\left[\tilde{D}^*(\omega,1)-\tilde{D}^*(-\omega,-1)\right] \tag{319}$$

$$= -\pi^2\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{1}{\left|\tilde{D}(\omega,1)\right|^2}f_0(p)f_0'(\omega)\delta(p-\omega). \tag{320}$$

We now have all elements for the determination of the right-hand side of the Lenard–Balescu equation (297). Inserting again the $1/N$ factor we obtain the following result which is valid only in one dimension

$$\frac{\partial f_0}{\partial t} = \frac{\pi^2}{N}\frac{\partial}{\partial p}\int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{|\tilde{D}(\omega,1)|^2}\delta(p-\omega)\left(f_0(\omega)\frac{\partial f_0(p)}{\partial p}-f_0(p)\frac{\partial f_0(\omega)}{\partial\omega}\right)=0. \tag{321}$$

We thus see that, in one dimension, the Lenard–Balescu operator vanishes: the diffusion term (first term in the r.h.s.) is exactly balanced by the friction term (second term in the r.h.s.). Consequently the collisional evolution is due to terms of higher order in $1/N$, and the Vlasov equation is valid for a longer time than previously expected. This remark was raised long ago in plasma physics [225].

The previous computation can be performed for a general (long-range) two-body potential in d dimensions. The result is the following general Lenard–Balescu equation [226,227]

$$\frac{\partial f(\mathbf{v},t)}{\partial t}=\frac{\pi(2\pi)^d}{m^2}\frac{\partial}{\partial \mathbf{v}}\cdot\int d\mathbf{v}_1 d\mathbf{k}\,\mathbf{k}\frac{\hat{u}(|\mathbf{k}|)^2}{|\tilde{D}(\mathbf{k},\mathbf{k}\cdot\mathbf{v}_1)|^2}\delta(\mathbf{k}\cdot(\mathbf{v}-\mathbf{v}_1))\left[\mathbf{k}\cdot\left(f(\mathbf{v}_1,t)\frac{\partial f(\mathbf{v},t)}{\partial \mathbf{v}}-f(\mathbf{v},t)\frac{\partial f(\mathbf{v}_1,t)}{\partial \mathbf{v}_1}\right)\right], \quad (322)$$

where the boldface variables are d-dimensional vectors, and where $\hat{u}(|\mathbf{k}|)$ is the real d-dimensional Fourier transform of the two-body potential; m is the mass of the particles and f is normalized to 1. The dielectric function $\tilde{D}(\omega,k)$ in this general case is given by

$$\tilde{D}(\omega,\mathbf{k})=1+\frac{(2\pi)^d}{m}\hat{u}(|\mathbf{k}|)\int d\mathbf{v}\frac{\mathbf{k}\cdot\frac{\partial f(\mathbf{v})}{\partial \mathbf{v}}}{\omega-\mathbf{k}\cdot\mathbf{v}}. \quad (323)$$

The $1/N$ "smallness" of the right-hand side of Eq. (322) is incorporated in the two-body potential. From this general expression we see that, while in $d=1$ the collisional term vanishes, it will instead be present for $d>1$. Thus, in a two-dimensional Coulombian plasma or for three-dimensional Newtonian interactions and in plasma physics, the Lenard–Balescu collisional term gives a contribution already at order $1/N$. Nicholson has derived its expression for general potentials for homogeneous cases [215]; Chavanis has continued this line of research [228,229] and obtained preliminary results for the inhomogeneous case [230,231]. Note however that collective effects are not taken into account in the latter works.

The Landau approximation, often used in plasma physics, corresponds to neglect collective effects in Eq. (322), which amounts to take $\tilde{D}(\omega,\mathbf{k})=1$, so that the structure of the Landau equation does not depend on the potential.

One notes that in a one-dimensional framework, the Lenard–Balescu and the Landau equations coincide: both of them reduce to the Vlasov equation. The collisional evolution is due to terms of higher order in $1/N$, implying that the collisional relaxation time scales as N^δ with $\delta>1$. Therefore, the system can remain frozen in a stationary solution of the Vlasov equation for a very long time, larger than N. Only non-trivial three-body correlations can induce further evolution of the system, as also remarked in Ref. [232] for 2D hydrodynamics. This is the reason of the very long relaxation time emphasized by numerical simulations [26].

5.2. Quasi-stationary states, diffusion and entropies

5.2.1. Numerical evidence of quasi-stationary states

For the HMF model, but this is also true for any one-dimensional long-range system, we have thus proven that Vlasov stable homogeneous distribution functions do not evolve on time scales of order smaller or equal to N.

The above result is an illuminating explanation of the numerical strong disagreement which was reported in Refs. [128, 233] between constant energy molecular dynamics simulations and canonical statistical mechanics calculations. For energies slightly below the second order phase transition energy (see Fig. 32), numerical simulations in the microcanonical ensemble show that the system is trapped for a long time, whose duration increases with the number N of particles, in a state far from that predicted by equilibrium statistical mechanics. Since the latter were initially derived using the canonical ensemble, these results had been believed, at first, to be the fingerprint of inequivalence between microcanonical and canonical ensembles. However, although the interaction is long-range, both ensembles lead to the same results for the HMF model, where only a second order phase transition occurs as shown in Section 4.4. More careful numerical experiments [26] have revealed the tendency of the simulation points, i.e., of the out-of-equilibrium state in which the system is trapped, to lie on the continuation to lower energies of the supercritical branch with zero magnetization of the caloric curve. These states have been called quasi-stationary states (QSS). More systematic simulations [26] have determined a $N^{1.7}$ scaling law for the duration of the QSS, at the end of which the system eventually evolves towards the Boltzmann–Gibbs equilibrium state. In Fig. 33, we display the time evolution of the magnetization, $m(t)$, with increasing particle number, showing the increase of the duration of the QSS: the power-law increase is evidenced by the choice of the logarithmic scale in the abscissa. Since this scaling law has been found when the system is initially prepared in a state that is a stable stationary solution of the Vlasov equation (in particular, for initial states homogeneous in θ), this numerical evidence is in agreement with the result derived above, that Vlasov stable homogeneous distribution functions do not evolve on time scales of order smaller or equal to N. Obviously, even if the initial state is Vlasov stable, finite N effects will eventually drive the system away from it and towards the Boltzmann–Gibbs equilibrium state.

If the initial condition is Vlasov unstable, a rapid evolution will take place. Simulations have also been performed in this case [233], showing that, after an initial transient, the systems remains trapped in other types of QSS, with different scaling laws in N for their duration. The existence of an infinite number of Vlasov stable distributions is actually the key point to explain the out-of-equilibrium QSS observed in the HMF dynamics. Let us show that the system evolves through other stable stationary states. In order to check the stationarity and the stability of an initial distribution $f_0(\theta,p)$, it is possible to study the temporal evolution of the magnetization, which is constant if the system is stable and stationary. Other possible macrovariables are the moments of the distribution function. It can be easily shown that any distribution of the form $f(\theta,p,0)=f_0(\theta,p)=F(p^2/2-m_x\cos\theta-m_y\sin\theta)\equiv f_0(e)$, with generic F, is a stationary solution of the

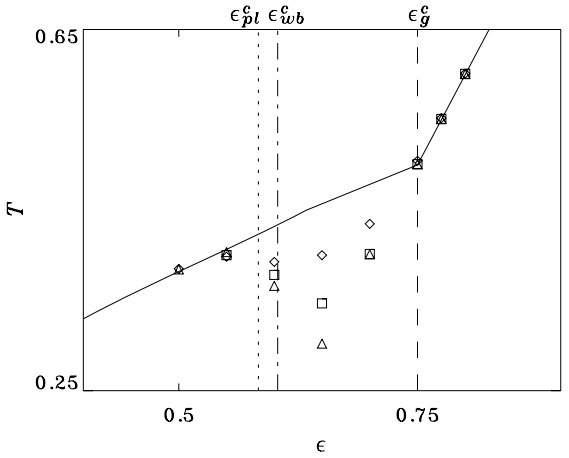


Fig. 32. Caloric curve of the HMF Hamiltonian (126). The solid line is the equilibrium result in both the canonical and the microcanonical ensemble. The second order phase transition is revealed by the kink at $\varepsilon^c_g = 3/4$. The three values of the energy indicated by the vertical lines are the stability thresholds for the homogeneous Gaussian (dashed), power-law of Eq. (295) with $\nu = 8$ (dash–dotted) and water-bag (dotted) initial momentum distribution. The Gaussian stability threshold coincides with the phase transition energy. The points are the results of constant energy (microcanonical) simulations for the Gaussian (losanges), the power-law (squares) and the water-bag (triangles). Simulations were performed with $N = 5000$.

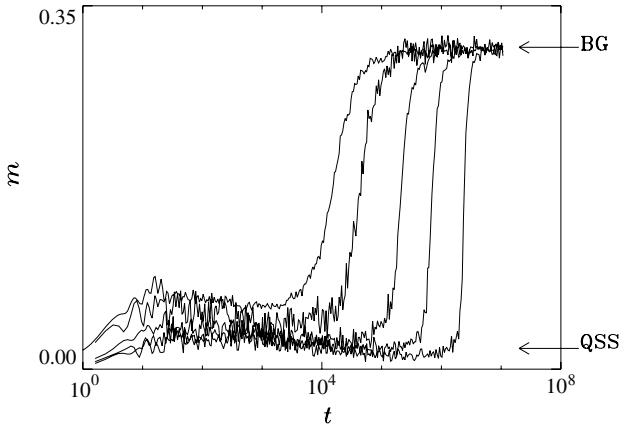
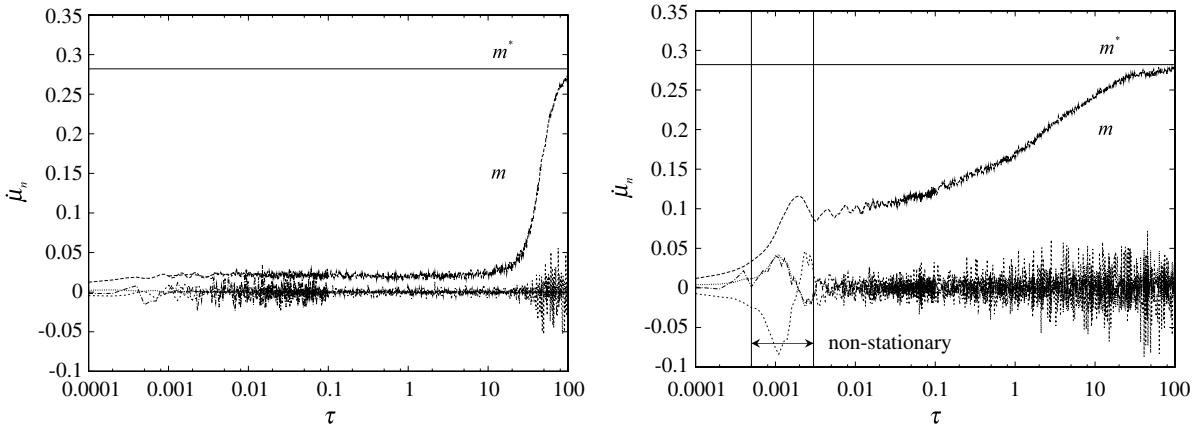


Fig. 33. Time evolution of the modulus of the magnetization $m(t)$ for different particle numbers: $N = 10^3, 2 \times 10^3, 5 \times 10^3, 10^4$ and 2×10^4 from left to right ($\varepsilon = 0.69$). In all cases, an average over several samples has been taken. Two values of the magnetization, indicated by horizontal arrows, can be identified in this figure: the upper one (labeled BG) corresponds to the expected equilibrium result for the magnetization, while the lower one, labelled QSS, represents the value of M in the quasi-stationary state.

Vlasov equation, provided m_x and m_y are determined self-consistently from Eq. (273). Obviously this does not yet mean that this solution is also stable. This latter observation has suggested to study numerically the stability of f_0 by checking the stationarity of the first few moments $\mu_n = \langle e^n \rangle_N$. Since the stationarity of the moments is a necessary condition for stability, the vanishing, for a long time lapse, of the time derivatives $\dot{\mu}_n = d\mu_n/dt$, for $n = 1, 2$ and 3, has been used as a numerical suggestion that the system is in a QSS and that the distribution $f_0(\theta, p)$ is a stable stationary solution of the Vlasov equation. On the contrary, large derivatives clearly indicate a non-stationary state.

Fig. 34 presents the temporal evolution of these quantities, together with the temporal evolution of the modulus m of the magnetization, for power-law ($\nu = 8$) and Gaussian initial distributions in the case of an energy ε in the interval $[\varepsilon^c_{\rm pl}, \varepsilon^c_{\rm g}]$. In the stable case (a), the stationarity holds throughout the computed time since one notices that the three quantities $\dot{\mu}_n$ have vanishingly small fluctuations around zero. On the contrary, in the unstable case (b), the system is first in an unstable stationary state ($\tau < 0.0005$), before becoming non-stationary ($0.0005 < \tau < 0.003$) and finally reaches stable stationary states ($\tau > 0.003$). Consequently the system evolves among different Vlasov stationary states. In the stable case (a), the magnetization m stays around zero before taking off around $\tau = 20$ to reach the equilibrium value m^*. In the unstable case, Fig. 34(b) shows that after experiencing unstable stationary and non-stationary states, the system presents a slow quasi-stationary evolution across the infinite number of stationary and stable Vlasov states. In Ref. [26], a careful numerical study has shown that this slow characteristic timescale associated to this final relaxation toward the Boltzmann–Gibbs equilibrium is proportional to $N^{1.7}$. However, this law might be dependent on the energy ε and on the initial distributions. The previous subsection allows however to claim that this relaxation timescale is larger than N at least. This is thus a very



(a) Distribution with power-law tails: Stable.

(b) Gaussian distribution: Unstable.

Fig. 34. Time evolution of the magnetization m. The time is rescaled as $\tau = t/N$. The quantities $\dot{\mu}_n$ ($n = 1, 2, 3$), which detect the stationarity, are also plotted (multiplied by a factor 100 for graphical purposes). The equilibrium value of the magnetization is m^*. Panel (a) corresponds to a stable homogeneous initial distribution with power-law tails, while panel (b) shows an unstable initial condition with Gaussian tails. In both panels, energy is $\varepsilon = 0.7$. All numerical curves are obtained by averaging over 20 initial conditions with $N = 10^4$.

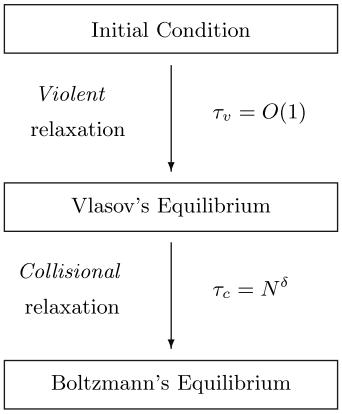


Fig. 35. Schematic description of the typical dynamical evolution of systems with long-range interactions. τ_v and τ_c are the *violent relaxation* and the *collisional relaxation* timescales, respectively.

slow process in comparison to the relaxation from an initially unstable state. Recently, this slow evolution of the HMF system through different stable stationary states of the Vlasov equation has been more systematically studied, with the aim to determine how the Vlasov stable solutions characterizing the system during the out-of-equilibrium dynamics can be parameterized [222,234]. It has been found that, starting from a homogeneous distribution given by a q-Gaussian with compact support, at an energy slightly below the second order phase transition energy, the evolution of the system during the QSS is well approximated by distribution functions of the same type, but with varying q, until the system heads towards Boltzmann–Gibbs equilibrium.

Let us stress that the above scenario is consistent with what happens generically for systems with long-range interactions [18,34,235]. In a first stage, called *violent relaxation*, the system goes from a generic initial condition, which is not necessarily Vlasov stable, towards a Vlasov stable state. This is a fast process happening usually on a fast timescale, *independent* of the number of particles. In a second stage, named *collisional relaxation*, finite N effects come into play and the Vlasov description is no more valid for the discrete systems. The timescale of this second process is strongly dependent on N. One generally considers that it is a power law N^δ. A typical example is the Chandrasekhar relaxation time scale for stellar systems, which is proportional to $N/\ln N$. This scenario of the typical evolution of long-range systems is summarized in Fig. 35.

It is important to remark that, recently, Caglioti and Rousset [236] rigorously proved that for a wide class of potentials, particles starting close to a Vlasov stable distribution remain close to it for times that scale at least like $N^{1/8}$: this result is consistent with the power law conjectured for collisional relaxation. Unfortunately, apart from a recent progress [237], very few rigorous results exist in the case of singular potentials, which would be of paramount importance for Coulomb and gravitational interactions.

Stronger divergences with system size in long-range systems are observed in connection with metastable states [161,238,239] where the relaxation time increases exponentially with N.

In summary, quasi-stationary states observed in the N-particle dynamics of the HMF model are nothing but Vlasov stable stationary states, which evolve because of *collisional*, finite N, effects. There is an *infinity* of Vlasov stable homogeneous (zero magnetization) states corresponding to different initial velocity distributions $f_0(t = 0, p)$, whose stability domain in energy are different. The q-Gaussians in momentum homogeneous distributions are Vlasov stable stationary states in a certain energy region where QSS are observed in the HMF model. However, they are not special in any respect, among an *infinity* of others. In the HMF model at finite N, all of them converge sooner or later to the Boltzmann–Gibbs equilibrium. However, the relaxation time is shown numerically to diverge with a power-law N^{δ}, with $\delta \simeq 1.7$ for the homogeneous water-bag state.

The time scale $\tau = t/N$ is thus the appropriate one to study momentum autocorrelation functions and diffusion in angle. We will consider such issues in the next subsection, where *weak* or *strong anomalous diffusion* for angles is predicted, both at equilibrium and for QSS.

5.2.2. Fokker–Planck equation for the stochastic process of a single particle

Let us now consider the relaxation properties of a test particle, initially with a momentum p_1 and an angle θ_1, immersed in a homogeneous background of N particles; the latter is consequently a thermal bath, or a reservoir, for the test particle. The description of the motion for a so-called test particle in a system with identical particles is a classical problem in kinetic theory. Initially the test particle is in a given microscopic state, while the other particles are distributed according to the distribution f_0 and the test particle is assumed not to affect the reservoir. The interaction with the fluid induces a complicated stochastic process.

The test particle distribution is initially not in equilibrium, and not necessarily close to it. However, it is natural to expect that the distribution of the test particle will eventually correspond to the distribution of the bath generated by all the other particles. How it evolves from the initial Dirac distribution $f_1(\theta, p, 0) = \delta(\theta - \theta_1)\delta(p - p_1)$ toward the equilibrium distribution is thus of high interest. We will show that the distribution is a solution of a Fokker–Planck equation that can be derived analytically. This equation describes the dynamical coupling with the fluctuations of the density of particles, which induce fluctuations in the potential: this is the origin of the underlying stochastic process.

We analyze therefore the relaxation properties of a test particle, indexed by 1, surrounded by a background system of $(N-1)$ particles with a homogeneous in angle distribution $f_0(p)$. The averaged potential $\langle v\rangle$ still vanishes for a homogeneous distribution, so that the particle only feels the fluctuations of the potential, that, according to Eq. (272), is given by $\delta v(\theta)/\sqrt{N}$. The potential felt by the test particle at the position $\theta_1(t)$ is therefore $\delta v(\theta)/\sqrt{N}$ computed at $\theta = \theta_1(t)$. We thus expect that the instantaneous force on the test particle will be of order $1/\sqrt{N}$. The equations of motion of the test particle are therefore

$$\frac{\mathrm{d}\theta_1}{\mathrm{d}t} = p_1 \quad \text{and} \quad \frac{\mathrm{d}p_1}{\mathrm{d}t} = -\frac{1}{\sqrt{N}}\left.\frac{\partial \delta v(\theta, t)}{\partial \theta}\right|_{\theta=\theta_1(t)}, \tag{324}$$

the integration of which leads to (omitting the index 1 for the sake of simplicity)

$$\theta(t) = \theta(0) + p(0)t - \frac{1}{\sqrt{N}}\int_0^t \mathrm{d}u_1 \int_0^{u_1} \mathrm{d}u_2 \frac{\partial \delta v}{\partial \theta}(\theta(u_2), u_2) \tag{325}$$

$$p(t) = p(0) - \frac{1}{\sqrt{N}}\int_0^t \mathrm{d}u \frac{\partial \delta v}{\partial \theta}(\theta(u), u), \tag{326}$$

where, again for simplicity, we have indicated directly inside the dependence of δv the test particle variable. The key point of this approach is that we do not limit the study to the usual ballistic approximation, in order to have an expansion exact at order $1/N$. Therefore it is of paramount importance here to treat accurately the essential *collective effects*.

By introducing iteratively the expression for the variable θ in the right-hand-side of Eq. (326) and by expanding the derivatives of the potential, one gets the result at order $1/N$ of the momentum dynamics

$$\begin{aligned} p(t) \;=\; & p(0) - \frac{1}{\sqrt{N}}\int_0^t \mathrm{d}u \frac{\partial \delta v}{\partial \theta}\,(\theta(0) + p(0)u, u) \\ & + \frac{1}{N}\int_0^t \mathrm{d}u \frac{\partial^2 \delta v}{\partial \theta^2}\,(\theta(0) + p(0)u, u) \int_0^u \mathrm{d}u_1 \int_0^{u_1} \mathrm{d}u_2 \frac{\partial \delta v}{\partial \theta}(\theta(0) + p(0)u_2, u_2). \end{aligned} \tag{327}$$

As the changes in the momentum are small (of order $1/\sqrt{N}$), the description of the momentum dynamics is well represented by a stochastic process governed by a Fokker–Planck equation [240]. If we denote by $f_1(p, t)$ the distribution function at time t of the test particle momentum, then the general form of this equation is

$$\frac{\partial f_1(p, t)}{\partial t} = -\frac{\partial}{\partial p}\left[A(p, t) f_1(p, t)\right] + \frac{1}{2}\frac{\partial^2}{\partial p^2}\left[B(p, t) f_1(p, t)\right], \tag{328}$$

with

$$A(p,t) = \lim_{\tau\to 0}\frac{1}{\tau}\langle (p(t+\tau)-p(t))\rangle_{p(t)=p} \tag{329}$$

$$B(p,t) = \lim_{\tau\to 0}\frac{1}{\tau}\langle (p(t+\tau)-p(t))^2\rangle_{p(t)=p}, \tag{330}$$

where the expectation values are conditioned by $p(t) = p$. This equation is therefore characterized by the time behavior of the first two moments, called Fokker–Planck coefficients. We approximate these two coefficients by an expression which is valid in the range of time t defined by $1 \ll t \ll N$. In this time range, using a generalization of formula (313) that takes into account that the initial coordinates of the test particles are given, it is possible to obtain (see Appendix E)

$$A(p,t) \sim \frac{1}{N}\left(\frac{\mathrm{d}D}{\mathrm{d}p}(p) + \frac{1}{f_0}\frac{\partial f_0}{\partial p}D(p)\right) \tag{331}$$

$$B(p,t) \sim \frac{2}{N}D(p), \tag{332}$$

where the diffusion coefficient is given by

$$D(p) = 2\,\mathrm{Re}\int_0^{+\infty}\mathrm{d}t\mathrm{e}^{\mathrm{i}pt}\left\langle\widehat{\delta v}(1,t)\widehat{\delta v}(-1,0)\right\rangle = \pi^2\frac{f_0(p)}{\left|\tilde{D}(p,1)\right|^2}. \tag{333}$$

Substituting (331) and (332) in the general form of the Fokker–Planck equation (328), one ends up with

$$\frac{\partial f_1(p,t)}{\partial t} = \frac{1}{N}\frac{\partial}{\partial p}\left[D(p)\left(\frac{\partial f_1(p,t)}{\partial p} - \frac{1}{f_0}\frac{\partial f_0}{\partial p}f_1(p,t)\right)\right]. \tag{334}$$

We thus recover what has been established in plasma physics (see e.g. Ref. [241]): the evolution of the velocity distribution $f_1(p,t)$ of the test particle is governed by a Fokker–Planck equation that takes a form similar to the Lenard–Balescu equation (321) provided that we replace the distribution $f_0(p,t)$ of the bath by $f_1(p,t)$. The integro-differential equation is thus transformed in the Fokker–Planck differential equation. Similar results in a higher dimension have been obtained later in Refs. [227,229].

For one-dimensional systems, it has been shown in Section 5.1.5 that the Lenard–Balescu collision term cancels out so that the distribution function does not evolve on a time scale of order N. Since, on the other hand, the Fokker–Planck equation (334) shows that the relaxation time of a test particle towards the distribution of the bath is of order N, this implies that we can assume that the distribution of the particles f_0 is stationary when one studies the relaxation of a test particle. This is true for any distribution function $f_0(p)$ that is a stable stationary solution of the Vlasov equation. This is not true in higher dimensions, except for the Maxwellian distribution.

Eq. (334) emphasizes that the momentum distribution of particle 1 evolves on timescales of order N. We will thus introduce the timescale $\tau = t/N$, so that the Fokker–Planck equation can be rewritten as

$$\frac{\partial f_1}{\partial \tau} = \frac{\partial}{\partial p}\left[D(p)\left(\frac{\partial f_1}{\partial p} - \frac{1}{f_0}\frac{\partial f_0}{\partial p}f_1\right)\right], \tag{335}$$

valid for times τ at least of order 1. We see that the time derivative of f_1 vanishes if the distribution function f_1 of the test particle is equal to the quasi-stationary distribution f_0 of the surrounding bath. We expect that f_1, governed by this Fokker–Planck equation, will converge to f_0 in a time τ of order 1: this means that the distribution function f_1 of the test particle converges towards the quasi-stationary distribution f_0 of the surrounding bath. Thus, it does not converge towards the equilibrium Gaussian distribution, in complete agreement with the result that f_0 is stationary for times scales of order N. This result is not valid in higher dimensions [227]. It is important to stress here that collective effects are taken into account through the presence of the dielectric response function $\tilde{D}(p,1)$ in the denominator of the diffusion coefficient (see Eq. (333)).

Analyzing the stochastic process of equilibrium fluctuations in the particular case of homogeneous Gaussian distribution, Bouchet [242] derived the diffusion coefficient of a test particle in a equilibrium bath. His result is recovered when considering a homogeneous Gaussian distribution (292) in expression (333) since one gets

$$D(p) = \pi^2\frac{\frac{1}{2\pi}\sqrt{\frac{\beta}{2\pi}}\mathrm{e}^{-\beta p^2/2}}{\left[1-\frac{\beta}{2}+\frac{1}{2}\beta^{3/2}p\mathrm{e}^{-\beta p^2/2}\int_0^{p\sqrt{\beta}}\mathrm{e}^{t^2}/2\mathrm{d}t\right]^2+\frac{1}{8}\pi\beta^3p^2\mathrm{e}^{-\beta p^2}}, \tag{336}$$

plotted in Fig. 36. It is important to stress that such an expression leads to a diffusion coefficient with Gaussian-like tails $D(p) \sim \sqrt{\pi\beta/8}\mathrm{e}^{-\beta p^2/2}$.

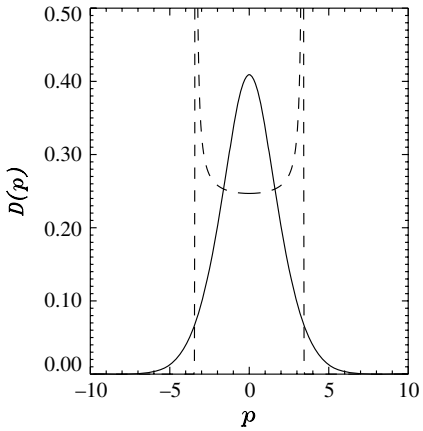


Fig. 36. Diffusion coefficient $D(p)$ in the case $\varepsilon = 2$ for a Boltzmann thermal bath (solid line) and a water bag distribution (dashed line).

The above general derivation [220] however allows one to study *any* arbitrary distribution. Let us carry on here the calculation of the diffusion coefficient for the Vlasov-stable water bag distribution (293). It has an interesting behavior, since the dielectric response $\tilde{D}(\omega, 1)$ has zeroes on the real axis, contrary to any even distributions *strictly* decreasing for positive values of the frequency ω. One gets

$$D(p) = \pi^2 \frac{\frac{1}{4\pi p_0}\left[\Theta(p+p_0) - \Theta(p-p_0)\right]}{\left[1 - \frac{1}{2}\frac{1}{p_0^2 - p^2}\right]^2 + \left[\frac{\pi}{4p_0}\left(\delta(p+p_0) - \delta(p-p_0)\right)\right]^2}, \tag{337}$$

which is also plotted in Fig. 36. One can obtain similar results for q-exponential distributions (294) (see Ref. [130]) or power-law tails distributions (295).

5.2.3. Long-range temporal correlations and diffusion

Since the Fokker–Planck equation for the single particle distribution function $f_1(p)$ has a variable diffusion coefficient, the relaxation towards the Boltzmann distribution can be slowed down, especially if the diffusion coefficient decreases rapidly with momentum. It is thus important to study this Fokker–Planck equation for different distribution functions of the bath. A similar study was earlier performed for 2D vortices in Ref. [232]. One consequence is that velocity correlation functions can decrease algebraically with time instead of exponentially, a behavior which might lead to anomalous diffusion as we will show below.

By introducing the appropriate change of variable $x = x(p)$, defined by $\mathrm{d}x/\mathrm{d}p = 1/\sqrt{D(p)}$, and the associated distribution function $\widehat{f}_1$, defined by $\widehat{f}_1(\tau, x)\mathrm{d}x = f_1(\tau, p)\mathrm{d}p$, one can map [220,243–246] the Fokker–Planck equation (335) to the constant diffusion coefficient Fokker–Planck equation

$$\frac{\partial \widehat{f}_1}{\partial \tau} = \frac{\partial}{\partial x}\left(\frac{\partial \widehat{f}_1}{\partial x} + \frac{\partial \psi}{\partial x}\widehat{f}_1\right), \tag{338}$$

where the potential $\psi(x)$ is given by

$$\psi(x) = -\ln\left(\sqrt{D(p(x))} f_0(p(x))\right). \tag{339}$$

Using the property $\tilde{D}(p, 1) \overset{|p|\to\infty}{\sim} 1$, that implies, by Eq. (333), that $D(p) \overset{|p|\to\infty}{\sim} \pi^2 f_0(p)$, one gets $\psi(x) \overset{x\to\pm\infty}{\sim} -\frac{3}{2}\ln f_0(p(x))$. From this, one derives that for many classes of distribution function f_0, the potential $\psi(x)$ is asymptotically equivalent to a logarithm. In fact, we have

$$\psi(x) \overset{x\to\pm\infty}{\sim} \alpha \ln|x|, \tag{340}$$

with $\alpha = 3$ if $f_0(p)$ decreases to zero more rapidly than algebraically for large p, and $\alpha < 3$ if $f_0(p)$ decreases to zero algebraically; more precisely, $\alpha = 3\nu/(2+\nu)$ if $f_0(p)$ decays at large p as $p^{-\nu}$. For weakly confining potentials $\psi(x)$, i.e., when $\nu < 1$ and thus $\alpha < 1$, Eq. (338) has a non-normalizable ground state. The example of the heat equation, which corresponds to $\psi(x) = 0$, describes a diffusive process leading to an asymptotic self-similar evolution. In such a case, the spectrum of the Fokker–Planck equation is purely continuous. By contrast, a strongly confining potential $\psi(x)$ (for instance, the Ornstein–Ulhenbeck process with a quadratic potential) would lead to exponentially decreasing distributions and autocorrelation functions, linked to the existence in the spectrum of a gap above the ground state. The logarithmic potential (340) is

a limiting case between the two behaviors. The normalizable ground state is unique and coincides with the bottom of the continuous spectrum. The absence of a gap forbids *a priori* any exponential relaxation. To illustrate this result, we evaluate the asymptotic behavior explicitly in two cases.

- Let us first consider distribution functions $f_0(p)$ with fast (more than algebraically) decreasing tails so that

$$f_0(p) \overset{|p|\to\infty}{\sim} C\exp(-\beta p^\delta), \tag{341}$$

which includes not only the Gaussian ($\delta = 2$) and exponential tails ($\delta = 1$), but also stretched-exponential ones with any arbitrary positive exponent δ. From the change of variable $\mathrm{d}x/\mathrm{d}p = 1/\sqrt{D(p)}$, asymptotic analysis leads to $p(x) \overset{|x|\to\infty}{\sim} (2\ln|x|/\beta)^{1/\delta}$ and to $\psi(x) \overset{x\to\pm\infty}{\sim} 3\ln|x|$. These estimates are sufficient (see Ref. [220] for details) to evaluate the long time behavior of the momentum autocorrelation function

$$\langle p(\tau)p(0)\rangle \overset{\tau\to+\infty}{\propto} \frac{(\ln\tau)^{2/\delta}}{\tau}, \tag{342}$$

which proves the existence of long-range temporal momentum autocorrelation for all values of δ, and therefore also in the case of Boltzmann equilibrium, $\delta = 2$.
- Let us now consider a distribution function $f_0(p)$ with algebraic tails

$$f_0(p) \overset{|p|\to\infty}{\sim} C|p|^{-\nu}. \tag{343}$$

In this case, one has $p(x) \overset{|x|\to\infty}{\sim} C'x^{2/(2+\nu)}$ and the asymptotic behavior (340) with, as we said, $\alpha = 3\nu/(2+\nu)$. We consider only cases where $\nu > 3$, to ensure that the second moment of the distribution f_0 (i.e., the average kinetic energy) does exist. For $\nu > 3$ one has that $9/5 < \alpha < 3$. The result for the momentum autocorrelation function is

$$\langle p(\tau)p(0)\rangle \overset{\tau\to+\infty}{\propto} \tau^{(3-\nu)/(2+\nu)}, \tag{344}$$

which characterizes an algebraic asymptotic behavior.

From the momenta autocorrelation, one usually derives the angle diffusion $\sigma_\theta^2(\tau) = \langle[\theta(\tau)-\theta(0)]^2\rangle = 2D_\theta\,\tau$ where D_θ is defined via the Kubo formula

$$D_\theta = \int_0^{+\infty} \mathrm{d}\tau\ \langle p(\tau)p(0)\rangle. \tag{345}$$

However, since the exponent $(3-\nu)/(2+\nu) = -1+5/(2+\nu)$ is larger than -1, the asymptotic result (344) shows that the integral (345) diverges. The asymptotic result (342) leads also to a divergent integral, although it is less singular. It is thus natural to expect anomalous diffusion for the angles. This extremely small anomaly (logarithmic) for distribution functions with Gaussian or stretched exponential tails induced difficulties to detect numerically anomalous diffusion [25,247].

Note that, generalizing the theory of Potapenko et al. [248], Chavanis and Lemou [249] studied how the structure and the progression of the distribution function tails, also called fronts, depends on the behavior of the diffusion coefficient for large velocities. They showed that the progression of the front is extremely slow (logarithmic) in that case so that the convergence towards the equilibrium state is peculiar.

Using the time rescaling $\tau = t/N$, which introduces a factor $1/N^2$, the angle diffusion $\sigma_\theta^2(\tau)$ can be rewritten as

$$\frac{\sigma_\theta^2(\tau)}{N^2} = \int_0^\tau \mathrm{d}\tau_1 \int_0^\tau \mathrm{d}\tau_2\ \langle p(\tau_1)p(\tau_2)\rangle \tag{346}$$

$$= 2\int_0^\tau \mathrm{d}s \int_0^{\tau-s} \mathrm{d}\tau_2\ \langle p(s+\tau_2)p(\tau_2)\rangle\,, \tag{347}$$

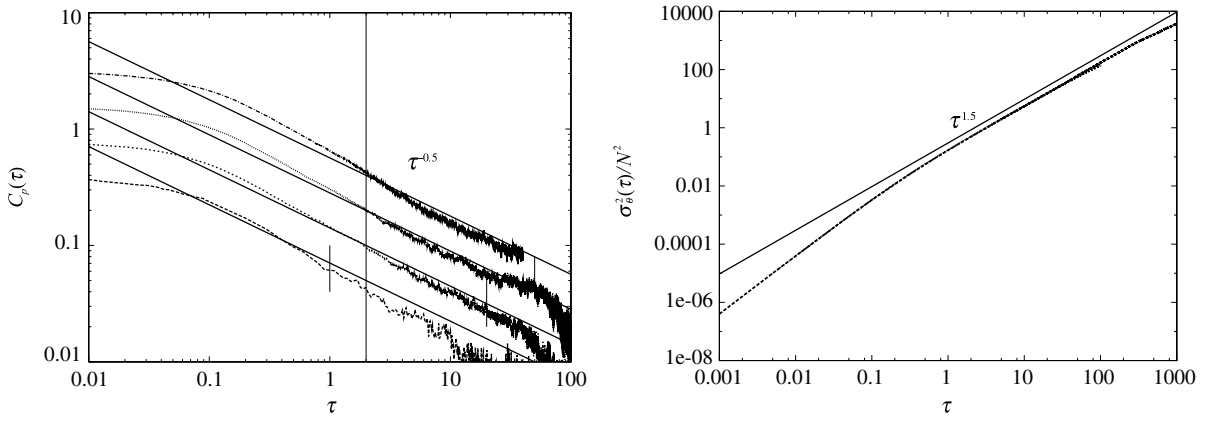
in which the new variable $s = \tau_1 - \tau_2$ has been introduced to take advantage of the division of the square domain into two isoscale triangles corresponding to $s > 0$ and $s < 0$. In the quasi-stationary states, the integrand $\langle p(s+\tau_2)p(\tau_2)\rangle$ does not depend on τ_2 (the QSS evolves on a time scale much larger than N) and hence diffusion can be simplified [247] as

$$\frac{\sigma_\theta^2(\tau)}{N^2} = 2\int_0^\tau \mathrm{d}s\,(\tau - s)\,\langle p(s)p(0)\rangle\,. \tag{348}$$

A distribution with power law tails (343) will therefore correspond to

$$\langle(\theta(\tau)-\theta(0))^2\rangle \overset{\tau\to+\infty}{\propto} \tau^{1+\frac{5}{2+\nu}}. \tag{349}$$

A comparison of this predicted [220] anomalous diffusion for angles with direct numerical computation of the HMF dynamics is a tough task because of the scaling with N of the time dependence of the autocorrelation function. However, it has been recently confirmed by numerical simulations in Ref. [223]. For the stable case ($\varepsilon = 0.7$) with power-law tails ($\nu = 8$), the



(a) Correlation.

(b) Anomalous diffusion.

Fig. 37. Check of the theoretical prediction for stable initial distributions with power-law tails, in the case $\varepsilon = 0.7$. Points are numerically obtained by averaging 20, 20, 10 and 5 realizations for $N = 10^3$, 10^4, 2×10^4 and 5×10^4 respectively. In panel (a), four curves represent the correlation functions of momenta, while the straight lines with the slope $-1/2$ represent the theoretical prediction. The curves and the lines are multiplied from the original vertical values by 2, 4 and 8 for $N = 10^4$, 2×10^4 and 5×10^4 for graphical purposes. Similarly, panel (b) presents the diffusion of angles, while the straight line with the slope $3/2$ is theoretically predicted. The four curves for the four different values of N are reported and almost collapse.

theory predicts that the correlation function decays algebraically with the exponent $-1/2$ (see Eq. (344)). Fig. 37(a) shows that the theoretical prediction agrees well with numerical computations. The expression (349) can be rewritten in that case as $\sigma^2_\theta(\tau) \sim \tau^{3/2}$: Fig. 37(b), in which the four curves for the four different values of N almost collapse, attests also the validity of this prediction. This is a clear example of anomalous diffusion.

For the stretched exponential distribution function (341), one ends up with

$$\langle(\theta(\tau) - \theta(0))^2\rangle \overset{\tau\to+\infty}{\propto} \tau(\ln\tau)^{2/\delta+1}. \tag{350}$$

As for power law tails, the diffusion is again anomalous, although with a logarithmically small anomaly. Consequently, the anomalous diffusion for angles also occurs for the Gaussian distribution which corresponds to the special case $\delta = 2$. This weak anomalous diffusion, i.e. normal diffusion with logarithmic corrections, has also been confirmed [223]. Since Gaussian distributions correspond to equilibrium distributions in the microcanonical and the canonical ensembles, it is important to realize that anomalous diffusion might thus be encountered for both equilibrium and out-of-equilibrium initial conditions. An analogous behavior has been observed for point vortices [34,246].

5.2.4. Lynden-Bell's entropy

In Section 5.2.1 we have discussed in detail, focusing on the HMF model, the two-stage relaxation process that is often observed in systems with long-range interactions. It has been indeed realized, beginning with a seminal paper by Hénon [121] on globular clusters, that the dynamical evolution is divided in two well separated phases. A first phase, called *dynamical mixing*, where an initial fast (violent) evolution leads to a *quasi-stationary state*, and a second phase, called *relaxation phase*, where "collisions" have the cumulative effect of driving the systems towards statistical equilibrium. If the number of particles is large, the two phases are well separated in time. It was also observed that the quasi-stationary state was strongly dependent on the initial condition [121]. In Section 5.1.4 we have discussed the existence of stationary stable and unstable one-particle distributions for the Vlasov equation and we have presented (Section 5.2.1) the interpretation of quasi-stationary states in terms of "attractive" Vlasov equilibria. A statistical approach that explains the existence of Vlasov equilibria has been proposed long ago by Lynden-Bell [23]. He begins by remarking that the Vlasov equation, which represents the evolution of an incompressible fluid, obeys the Liouville theorem in six dimensions. We will here restrict to a two-dimensional phase space for simplicity. This implies that the "mass" of phase elements between f and $f + \mathrm{d}f$ is conserved (remind that f is the one-particle distribution function). If we discretize the one-particle distribution function into a set of k quantized levels η_i, $i = 1, \ldots, k$, this means that the area corresponding to the i-th level $m(\eta_i) = \int \mathrm{d}\theta\mathrm{d}p\delta(f(\theta, p, t) - \eta_i)$ is a constant of the motion. In the limit of a continuous distribution of levels, one obtains an infinity of conservation laws. It is easy to prove that this implies that any functional of the form $\int \mathrm{d}\theta\mathrm{d}p\, C(f)$ is conserved: these functionals are called Casimirs. For instance, Gibbs entropy is a particular Casimir: $S_{Gibbs} = -\int \mathrm{d}\theta\mathrm{d}p f \ln f$. This specifically implies that, in terms of the *fine grained* one-particle distribution, Gibbs entropy cannot increase. Let us then define a *coarse grained* distribution

$$\bar{f}(\theta, p) = \sum_{i=1}^{k} \rho(\theta, p, \eta_i)\, \eta_i, \tag{351}$$

where $\rho(\theta, p, \eta_i)\mathrm{d}\theta\mathrm{d}p$ is the probability of finding level η_i in the phase-space macrocell $D_{macro} = [\theta, \theta + \mathrm{d}\theta] \times [p, p + \mathrm{d}p]$. Obtaining this probability directly from the microscopic dynamics is an extremely difficult problem, and its solution would

constitute a significant step forward in the understanding of the nature of quasi-stationary states. Lynden-Bell's proposal is to evaluate $\rho(\theta, p, \eta_i)$ by using a sort of "Boltzmann principle" [6,56,57,151,152]. The evaluation of $\rho(\theta, p, \eta_i)$ can be done for any number of levels k, but to make the calculation simpler, let us approximate the one-particle distribution with only two levels $\eta_1 = 0$ and $\eta_2 = f_0$ (the case with many levels has been treated in Refs. [4,23]). Alternatively, one can consider a fine-grained distribution function which, at $t = 0$, has only two levels: these are called *water-bags* in astrophysics and plasma physics. The exact time evolution of the water-bag is such that the shape of the area occupied by the Vlasov "fluid" with level f_0 is deformed, stretched and folded due to the hyperbolicity originated by the nonlinear dynamics of the Vlasov equation (see e.g. Fig. 2 in Ref. [23]). However, due to Liouville's theorem, this area is conserved. Let us divide each macrocell D_{macro} into ν microcells of volume ω and consider a microscopic configuration in which the i-th macrocell is occupied by n_i microcells with level f_0 and $\nu - n_i$ with level zero. The total number of occupied microcells is $\mathcal{N}$, such that the total mass is $m = \mathcal{N}\omega f_0$. This latter is also equal to the normalization of the fine grained distribution $m = \int \mathrm{d}\theta \mathrm{d}p f(\theta, p)$. The $\mathcal{N}$ occupied microcells are first placed into macrocells. There are $\mathcal{N}!/\prod_i n_i!$ ways to do this. Within the i-th cell, one can distribute the first of the n_i occupied microcells in ν ways, the second in $\nu - 1$ and so on. The number of ways of assigning the n_i occupied microcells is thus $\nu!/(\nu - n_i)!$. Then, the total number of microstates compatible with the macrostate where n_i microcells are occupied in macrocell i is given by the product of these two factors

$$W(\{n_i\}) = \frac{\mathcal{N}!}{\prod_i n_i!} \times \prod_i \frac{\nu!}{(\nu - n_i)!}. \tag{352}$$

The first factor is calculated exactly as for a Boltzmann gas [10], because the occupied microcells are *distinguishable*, while the second factor recalls Fermi–Dirac statistics and derives from an *exclusion principle*, which is a consequence of fluid incompressibility: one microcell cannot be occupied more than once by a fluid element of level f_0. Apart from this latter constraint, fluid microcells are let to distribute freely among the different macrocells: this corresponds to making an assumption of ergodicity. This does not happen for the true Vlasov dynamics and is sometimes referred to as the hypothesis of *efficient mixing*. Indeed, it has been found that dynamical effects can hinder mixing [205,250–255]. Using Stirling's approximation and expressing n_i in terms of the average probability to find level f_0 in cell i, $\rho_i(f_0) = n_i/\nu$, one obtains, neglecting an additive constant,

$$\ln W = \nu \sum_i \rho_i \ln \rho_i + (1 - \rho_i) \ln(1 - \rho_i), \tag{353}$$

which can also be rewritten in terms of the coarse grained distribution function, $\rho_i = \bar{f}_i/f_0$. Taking the continuum limit $\sum_i \to \int \mathrm{d}\theta \mathrm{d}p/(\omega\nu)$, one finally gets

$$s_{LB}[\bar{f}] = -\int \frac{1}{\omega} \mathrm{d}\theta \mathrm{d}p \left[\frac{\bar{f}}{f_0} \ln \frac{\bar{f}}{f_0} + \left(1 - \frac{\bar{f}}{f_0}\right) \ln \left(1 - \frac{\bar{f}}{f_0}\right) \right]. \tag{354}$$

Following the standard procedure, inspired by large deviation theory [99], one then maximizes s_{LB} subject to the constraint of conserving energy E, mass m and other global invariants like momentum (or angular momentum for higher dimensions).

Historically, the first applications of Lynden-Bell's ideas encountered both confirmations and failures [256–259], although the crucial point of necessarily performing simulations with a large number of particles was never clearly addressed, due to computer time limitations. Very recently, a careful analysis of the one-dimensional self-gravitating sheet model has shown that Lynden-Bell statistics applies for initial data with a virial ratio close to unity [260]. A statistical theory similar to Lynden-Bell's was independently developed for the Euler equation [7,151,261] and, later on, the deep analogy between the Vlasov–Poisson system and the Euler equation was for the first time clearly stressed in Ref. [24]. This approach was then used in the context of mean-field models in Refs. [126,262], with an application to the Colson–Bonifacio model of the Free Electron Laser (see Section 4.5.6). As already discussed in Section 5.2.1, the presence of quasi-stationary states for the HMF model was recognized and characterized numerically in Ref. [26]. Lynden-Bell's theory was then shown to predict the main features of the one-particle distribution function of quasi-stationary states of the HMF model in Ref. [263]. Let us finish this subsection by discussing in some detail this latter result, which clearly shows the power of Lynden-Bell's approach.

Let us first recall the Vlasov equation for the HMF model

$$\frac{\partial f}{\partial t} + p \frac{\partial f}{\partial \theta} - \frac{\mathrm{d}V}{\mathrm{d}\theta} \frac{\partial f}{\partial p} = 0, \tag{355}$$

where $f(\theta, p, t)$ is the *fine grained* one-particle distribution function and

$$V(\theta)[f] = 1 - M_x[f] \cos(\theta) - M_y[f] \sin(\theta), \tag{356}$$

$$M_x[f] = \int_{-\pi}^{+\pi} \int_{-\infty}^{+\infty} f(\theta, p, t) \cos\theta \mathrm{d}\theta \mathrm{d}p, \tag{357}$$

$$M_y[f] = \int_{-\pi}^{+\pi} \int_{-\infty}^{+\infty} f(\theta, p, t) \sin\theta \mathrm{d}\theta \mathrm{d}p. \tag{358}$$

The globally conserved quantities are energy

$$h[f] = \iint \frac{p^2}{2} f(\theta, p, t) \mathrm{d}\theta \mathrm{d}p - \frac{M_x^2 + M_y^2 - 1}{2}, \tag{359}$$

and momentum

$$P[f] = \iint p f(\theta, p, t) \mathrm{d}\theta \mathrm{d}p. \tag{360}$$

As for the initial distribution, we consider a *water bag* with rectangular shape in the (θ, p) plane. The distribution f takes only two distinct values, namely $f_0 = 1/(4\Delta\theta\Delta p)$, if the angles (velocities) lie within an interval centered around zero and of half-width $\Delta\theta$ (Δp), and zero otherwise ("mass" m is normalized to one and momentum $P[f]$ is zero). There is a one to one relation between the parameters $\Delta\theta$ and Δp and the initial values of magnetization and energy

$$M_0 = \frac{\sin(\Delta\theta)}{\Delta\theta}, \qquad e = \frac{(\Delta p)^2}{6} + \frac{1 - (M_0)^2}{2}. \tag{361}$$

While $h[f] = e$ and $P[f] = 0$ are constants of the motion, magnetization $M = \sqrt{M_x^2 + M_y^2}$ evolves with time. Lynden-Bell's maximum entropy principle is then defined by the following constrained variational problem

$$s(e) = \max_{\bar{f}} \left(s(\bar{f}) | h(\bar{f}) = e; P(\bar{f}) = 0; \int \mathrm{d}\theta \mathrm{d}p \bar{f} = 1 \right). \tag{362}$$

The problem is solved by introducing three Lagrange multipliers β/f_0, λ/f_0 and μ/f_0 for energy, momentum and mass normalization. This leads to the following analytical form of the distribution

$$\bar{f}(\theta, p) = \frac{f_0}{1 + \exp\left[\beta(p^2/2 - M_y[\bar{f}] \sin\theta - M_x[\bar{f}] \cos\theta) + \lambda p + \mu\right]}. \tag{363}$$

This distribution differs from the Boltzmann–Gibbs one because of the "fermionic" denominator. Inserting expression (363) into the energy, momentum and normalization constraints and using the definition of the magnetization, it can be straightforwardly shown that the momentum multiplier vanishes, $\lambda = 0$. Moreover, defining $x = \mathrm{e}^{-\mu}$ and $\mathbf{m} = (\cos\theta, \sin\theta)$, yields the following system of implicit equations in the unknowns β, x, M_x and M_y

$$f_0 \frac{x}{\sqrt{\beta}} \int \mathrm{d}\theta \mathrm{e}^{\beta \mathbf{M}\cdot\mathbf{m}} F_0\left(x \mathrm{e}^{\beta \mathbf{M}\cdot\mathbf{m}}\right) = 1 \tag{364}$$

$$f_0 \frac{x}{2\beta^{3/2}} \int \mathrm{d}\theta \mathrm{e}^{\beta \mathbf{M}\cdot\mathbf{m}} F_2\left(x \mathrm{e}^{\beta \mathbf{M}\cdot\mathbf{m}}\right) = e + \frac{M^2 - 1}{2} \tag{365}$$

$$f_0 \frac{x}{\sqrt{\beta}} \int \mathrm{d}\theta \cos\theta \, \mathrm{e}^{\beta \mathbf{M}\cdot\mathbf{m}} F_0\left(x \mathrm{e}^{\beta \mathbf{M}\cdot\mathbf{m}}\right) = M_x \tag{366}$$

$$f_0 \frac{x}{\sqrt{\beta}} \int \mathrm{d}\theta \sin\theta \mathrm{e}^{\beta \mathbf{M}\cdot\mathbf{m}} F_0\left(x \mathrm{e}^{\beta \mathbf{M}\cdot\mathbf{m}}\right) = M_y \tag{367}$$

with $F_0(y) = \int \exp(-v^2/2)/(1 + y\exp(-v^2/2))\mathrm{d}v$ and $F_2(y) = \int v^2 \exp(-v^2/2)/(1 + y\exp(-v^2/2))\mathrm{d}v$. This system of equations can be solved numerically and, given the parameters that fix the initial conditions, i.e. energy, momentum and M_0, univocally determines the values of the Lagrange multipliers β and μ, the values of M_x and M_y where Lynden-Bell's entropy is extremal and, finally, the distribution $\bar{f}$ in formula (363) itself, *with no adjustable parameter*. For $e = 0.69$, the maximum entropy state has zero magnetization for $M_0 < M_{crit} = 0.897$, a value at which Lynden-Bell's theory predicts a second order phase transitions [251,263,264]. Interpreting $\bar{f}$ as the distribution in the quasi-stationary state (QSS), we obtain in this case

$$\bar{f} = f_{QSS}(p) = \frac{f_0}{1 + \exp\left[\beta p^2/2 + \mu\right]}, \tag{368}$$

with β and μ to be determined from the knowledge of M_0. Velocity profiles predicted by (368) are displayed in Fig. 38 for different values of the initial magnetization.

Although not a single free parameter is used, one finds an excellent qualitative agreement. The presence of two symmetric bumps in the velocity distributions is not predicted by Lynden-Bell's theory and is a consequence of a collective phenomenon which leads to the formation of two *clusters* in the (θ, p) plane. This is shown by the direct simulation of the Vlasov equation (355) presented in Fig. 39. The bumps represent an intrinsic peculiarity of QSS and have been characterized dynamically in Ref. [252].

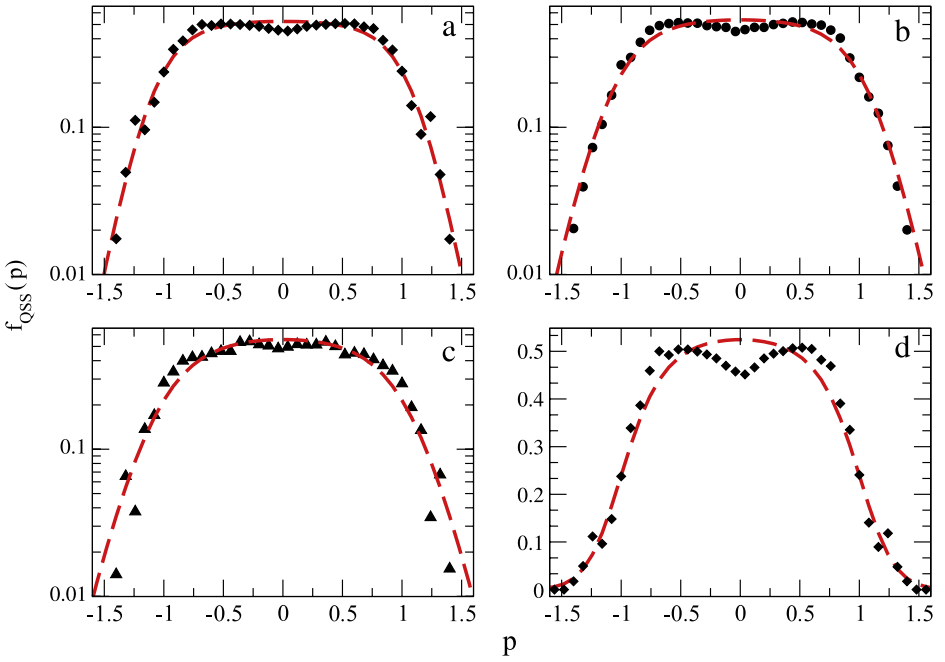


Fig. 38. Velocity distribution functions in the quasi-stationary state for the HMF model with $e = 0.69$ and different values of M_0. Symbols refer to numerical simulations, while dashed solid lines stand for the theoretical profile (368). Panels (a), (b) and (c) present the three cases $M_0 = 0.3$, $M_0 = 0.5$ and $M_0 = 0.7$ in lin-log scale, while panel (d) shows the case $M_0 = 0.3$ in lin–lin scale. The numerical curves are computed from one single realization with $N = 10^7$ at time $t = 100$.

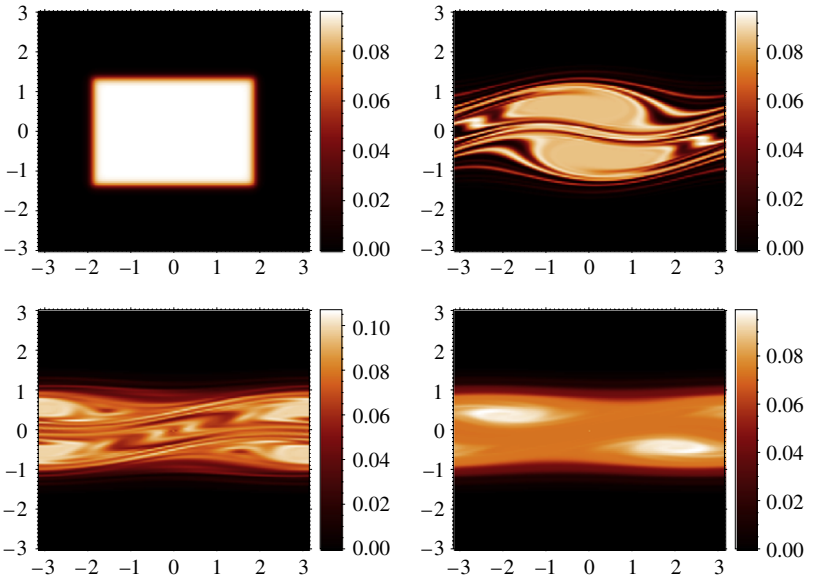


Fig. 39. Simulation of the Vlasov equation (355) that starts from a *water-bag* initial condition with $e = 0.69$ and $M_0 = 0.5$. The final snapshot (lower right panel) is the quasi-stationary state.

More recently, other authors have obtained similar encouraging results for Lynden-Bell's theory in a model of non-neutral plasma [253] and for radially symmetric solutions of a self-gravitating system [254]. They also find that, when dynamical effects lead to collective oscillations of the mean-field, Lynden-Bell's theory shows a disagreement with numerical data and propose a modification of the theory. Work along this line is in progress.

The theoretical approach proposed by Lynden-Bell [23] allows to predict the presence of a phase transition line in the (M_0, e) and in the (f_0, e) control parameter planes [251,264]. At equilibrium the phase transition does not depend on the choice of M_0 and is located at $e = 3/4$. On the contrary, in the out-of-equilibrium QSS the phase transition line is located at

smaller energies joining the transition points $(0, 7/12)$ [26] and $(1, 3/4)$ in the (M_0, e) plane. Moreover the phase transition changes from first to second order. These latter results are resumed in Ref. [265].

The Lynden-Bell approach turns out to be a good way to attack the problem of quasi-stationary states. It gives predictions for both averages and distributions functions which compare quite well with numerical simulations. When the approach fails to describe detailed features, they are viable ways of modifying it taking into account dynamical properties.

6. Generalization to non-mean-field models

Although substantial progress has been done recently in the understanding of long-range interactions, the hardest questions about physical systems where the interactions weakly decay with the distance, and also shows singularities at short distances, remain open (see the introductory Section 3). In Section 4, we have shown how simple mean-field models, that are explicitly solvable in both the canonical and the microcanonical ensemble, already show several features previously encountered in gravitational systems, e.g. negative specific heat. These simple models also display peculiar out-of-equilibrium dynamical effects, like the presence of quasi-stationary states, which resemble those found for realistic systems, see Section 5.

In this section, we make a little step in the direction of the study of Hamiltonians that are not fully mean-field. In Section 6.1 we introduce, for both a mean-field Ising and an XY model, a nearest-neighbor interaction term. We show that the phase diagram in canonical and microcanonical ensembles preserves the features of non-equivalence found for pure mean-field Hamiltonians. In particular, we mention the non coincidence of the microcanonical and canonical tricritical points and the presence of negative specific heat. In Section 6.2, we consider an Ising model in one dimension with weakly decaying coupling and a modification of the HMF model which includes a coupling with the same properties. For both models, we show that the mean-field properties extend to the weakly decaying case and that one can obtain analytically the free energy and the entropy. Since the corresponding mean-field limit of both models shows ensemble equivalence, this property persists also for such models. In Section 6.2.3, we show the analytical solution in both the canonical and microcanical ensembles of a self-gravitating system of particles moving in one dimension. The phase diagram is different in the two ensembles and shows all the features of ensemble inequivalence. In an appropriate limit, the model reduces to the HMF model and therefore ensemble inequivalence can be turned off by varying a parameter. Finally, in Section 6.2.4, we conclude with a summary of a very recent result concerning a spin system with dipolar interactions whose Hamiltonian can be reduced in appropriate limits to that of the XY model with nearest neighbor and mean-field interactions presented in Section 6.1. This result opens the possibility to verify experimentally some of the striking and counterintuitive features of long-range interactions.

6.1. Systems with short- and long-range interactions: The transfer integral method

In this subsection, we will present both the canonical and the microcanonical solutions of an Ising model and of an XY model with nearest neighbor and mean-field interactions. We will obtain the phase diagram in both ensembles. All the features of mean-field models discussed in Section 4 will appear again in this context: ensemble inequivalence, negative specific heat, temperature jumps, ergodicity breaking. This shows that the addition of a short-range term to a long-range Hamiltonian does not remove the interesting behaviors described before. For what concerns methodological aspects, the use of transfer integral combined with the Hubbard–Stratonovich transformations and the min–max method allows us to treat both models with discrete and continuous variable. The models we will consider are defined on a one-dimensional lattice (indeed a ring because of the periodic boundary conditions), but in principle there is no reason that forbids extending the calculation to strip and bar geometries.

6.1.1. Ising model

6.1.1.1. Introduction. Nagle introduced an interesting Ising model combining long- with short-range interactions [266]. Further elaborations of the model were proposed by Khardar [267,268]. The Hamiltonian is

$$H_N = -\frac{1}{2N}\left(\sum_{i=1}^{N} S_i\right)^2 - \frac{K}{2}\sum_{i=1}^{N}(S_i S_{i+1} - 1)\,, \tag{369}$$

where $S_i = \pm 1$. In this one-dimensional spin chain, the first term has an infinite range and is the typical one of the Curie–Weiss Hamiltonian (1). This term is responsible for the non-additive properties of the model, see Section 2.1. Note that the prefactor J present in Hamiltonian (1) has been set to one by an appropriate renormalization of the energy. There is no loss of generality, since only the ferromagnetic case $J > 0$ will be considered. On the contrary, the second term of (369) corresponds to an interaction between nearest neighbors along a one-dimensional lattice (periodic boundary conditions are chosen). The coupling constant K might be either positive or negative.

The ferromagnetic state with all spins up ($S_i = 1, \forall i$) or down ($S_i = -1, \forall i$) has a negative energy $E_F = -N/2$. For the antiferromagnetic state with alternate signs of nearest neighbor spins, the first term of (369) gives a vanishing contribution

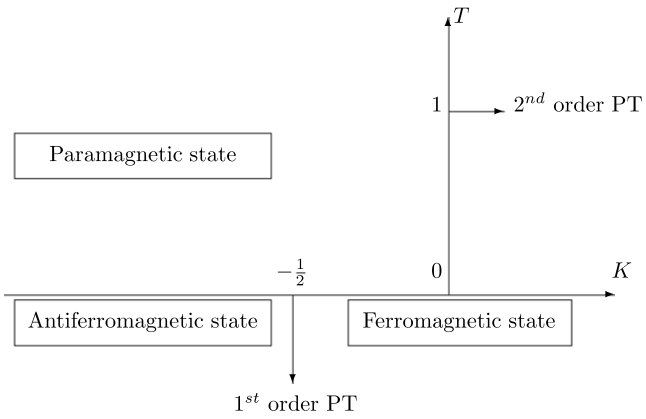


Fig. 40. Elementary features of the phase diagram of the short plus long-range Ising model showing the phase transitions on the temperature T and local coupling K axis, respectively.

to the energy and the energy is $E_A = KN$. At $T = 0$, one can determine if the model has a phase transition at some value of K by comparing the energy of the ferromagnetic state with that of the antiferromagnetic one, since only the energy term of the free energy matters. Hence, by imposing $E_A = E_F$ one gets the phase transition value $K_t = -1/2$ at which a discontinuity of the order parameter is found, from $m = 0$ to $m = 1$. Therefore the transition is first order.

For non-zero temperatures, one has to take into account the entropic term of the free energy, which measures disorder. When the coupling constant K vanishes, one fully recovers the Curie–Weiss Hamiltonian (1), which exhibits a *second* order phase transition at $T = 1$. One therefore expects that the (T, K) phase diagram displays a transition line which is first order at low T and second order at high T (see Fig. 40). The transition line separates a ferromagnetic from a paramagnetic state, although exactly at $T = 0$ and $K < -1/2$ the state is antiferromagnetic. Let us now determine analytically this transition line in both the canonical and the microcanonical ensemble.

6.1.1.2. The solution in the canonical ensemble. The phase diagram has been studied in the canonical ensemble by Nagle [266] and Khardar [267,268]. The partition function is

$$Z(\beta, N) = \sum_{\{S_1,\ldots,S_N\}} \mathrm{e}^{-\beta H} = \sum_{\{S_1,\ldots,S_N\}} \exp\left[\frac{\beta}{2N}\left(\sum_{i=1}^{N} S_i\right)^2 + \frac{\beta K}{2}\sum_{i=1}^{N}(S_i S_{i+1} - 1)\right]. \tag{370}$$

To get rid of the quadratic term, we use the Hubbard–Stratonovich transformation

$$\mathrm{e}^{\frac{\beta}{2N}\left(\sum_{i=1}^{N} S_i\right)^2} = \sqrt{\frac{\beta N}{2\pi}}\int_{-\infty}^{+\infty} \mathrm{d}x\, \mathrm{e}^{-\frac{\beta N}{2}x^2 + \beta x\sum_{i=1}^{N} S_i}, \tag{371}$$

so that the partition function (370) can be rewritten as

$$Z(\beta, N) = \sqrt{\frac{\beta N}{2\pi}}\int_{-\infty}^{+\infty} \mathrm{d}x\, \mathrm{e}^{-\frac{\beta N}{2}x^2} \sum_{\{S_1,\ldots,S_N\}}\left[\mathrm{e}^{\beta x\sum_{i=1}^{N} S_i + \frac{\beta K}{2}\sum_{i=1}^{N}(S_i S_{i+1}-1)}\right] \tag{372}$$

$$= \sqrt{\frac{\beta N}{2\pi}}\int_{-\infty}^{+\infty} \mathrm{d}x \mathrm{e}^{-N\beta\tilde{f}(\beta,x)}. \tag{373}$$

The free energy can be written as

$$\tilde{f}(\beta, x) = \frac{1}{2}x^2 + f_0(\beta, x), \tag{374}$$

where $f_0(\beta, x)$ is the free energy of the nearest-neighbor Ising model with an external field x. Such an expression can be easily derived using the transfer matrix [10,267,269]. By an easy calculation, we find $f_0(\beta, x) = -\ln(\lambda_+^N + \lambda_-^N)/(\beta N)$ where the two eigenvalues of the transfer matrix are

$$\lambda_\pm = \mathrm{e}^{\beta K/2}\cosh(\beta x) \pm \sqrt{\mathrm{e}^{\beta K}\sinh^2(\beta x) + \mathrm{e}^{-\beta K}}. \tag{375}$$

As $\lambda_+ > \lambda_-$ for all values of x, only the larger eigenvalue λ_+ is relevant in the limit $N \to \infty$. One thus finally gets

$$\tilde{\phi}(\beta, x) = \beta\tilde{f}(\beta, x) = \frac{\beta}{2}x^2 - \ln\left[\mathrm{e}^{\beta K/2}\cosh(\beta x) + \sqrt{\mathrm{e}^{\beta K}\sinh^2(\beta x) + \mathrm{e}^{-\beta K}}\right], \tag{376}$$

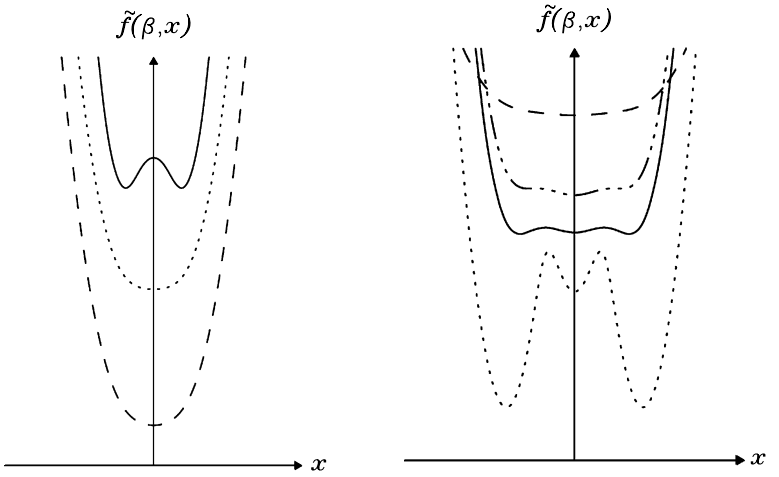


Fig. 41. $\tilde{f}(\beta, x)$ for different values of the inverse temperature. Left panel presents $\beta = 1.1$ (dashed line), $\beta_c \simeq 1.4$ (dotted), $\beta = 2.5$ (solid) when $K = -0.25$: a *second* order phase transition. Right panel shows the case $K = -0.4$ when $\beta = 10$ (dotted), $\beta_t \simeq 2.4$ (solid), $\beta = 2.35$ (dash-triple dot), $\beta = 2$ (dashed): a *first* order phase transition. Note that the different curves have been vertically shifted for readability purposes.

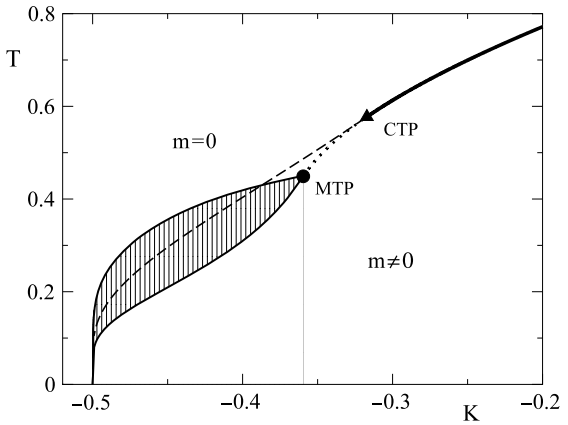


Fig. 42. The canonical and microcanonical (K, T) phase diagram. In the canonical ensemble, the large K transition is continuous (bold solid line) down to the tricritical point CTP where it becomes first order (dashed line). In the microcanonical ensemble the continuous transition coincides with the canonical one at large K (bold line). It persists at lower K (dotted line) down to the tricritical point MTP where it turns first order, with a branching of the transition line (solid lines). The shaded area is not accessible in the microcanonical ensemble.

which is shown in Fig. 41 for different values of the inverse temperature β and for two values of the nearest-neighbor coupling K. In the large N-limit, the application of the saddle point method to Eq. (373) finally leads to the free energy, which is obtained by taking the value of x which minimizes $\tilde{\phi}(\beta, x)$ in formula (376).

From the knowledge of the free energy, as anticipated, one gets either a second or a first order phase transition depending on the value of the coupling constant K. As usual, the expansion of $\tilde{f}(\beta, x)$ in power of x is the appropriate procedure to define the critical lines and points. One gets here

$$\tilde{f}(\beta, x) = -\ln 2\cosh\frac{\beta K}{2} + \frac{\beta}{2}x^2\left(1-\beta e^{\beta K}\right) + \frac{\beta^4}{24}e^{\beta K}\left(3e^{2\beta K}-1\right)x^4 + \mathcal{O}(x^6). \tag{377}$$

The critical point of the second order transition is obtained for each K by computing the value β_c at which the quadratic term of the expansion (377) vanishes provided the coefficient of the fourth order term is positive, obtaining $\beta_c = \exp(-\beta_c K)$. When the fourth order coefficient also vanishes, $3\exp(2\beta K) = 1$ one gets the canonical tricritical point (CTP) $K_{CTP} = -\ln 3/(2\sqrt{3}) \simeq -0.317$. The first order line is obtained numerically by requiring that $f(\beta, 0) = f(\beta, x^*)$, where x^* is the further local minimum of f. Fig. 42 represents the phase diagram in both the canonical and the microcanonical ensembles. The features of this phase diagram are very close to those of the BEG model (see Fig. 14) and of the generalized HMF (see Fig. 23). We will comment them below, after having obtained the solution in the microcanonical ensemble [161].

6.1.1.3. The solution in the microcanonical ensemble. In the microcanonical ensemble, a simple counting method has been recently proposed [161]. The magnetization $M = \sum_{i=1}^{N} S_i$ can be rewritten as $M = N_+ - N_-$ by introducing the number of

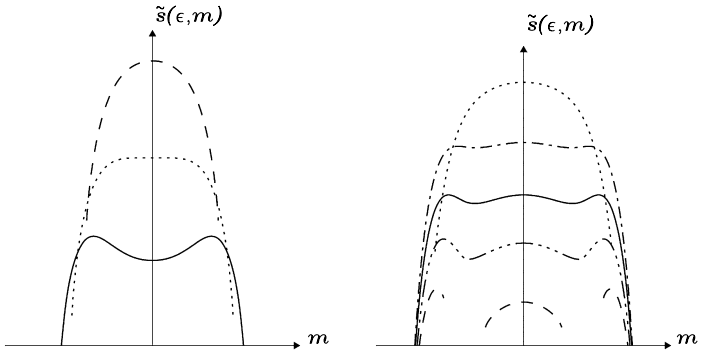


Fig. 43. $\tilde{s}(\varepsilon, m)$ for different values of the energy. Left panel presents $\varepsilon = -0.1$ (dashed line), $\varepsilon_c \simeq -0.15$ (dotted), $\varepsilon = -0.2$ (solid) when $K = -0.25$: a *second* order phase transition. Right panel shows the case $K = -0.4$ when $\varepsilon = -0.25$ (dotted), $\varepsilon = -0.305$ (dash–dotted), $\varepsilon_t = -0.3138$ (solid), $\varepsilon = -0.32$ (dash-triple dot), $\varepsilon = -0.33$ (dashed): a *first* order phase transition. The gaps present in the lower dashed curve are related to ergodicity breaking (see text). Note that the different curves have been vertically shifted for readability purposes.

spins up, N_+, and of spins down, N_-. The first term of the Hamiltonian (369) can be straightforwardly rewritten as $-M^2/(2N)$. As two identical neighboring spins would not contribute to the second term of Hamiltonian (369) ($S_iS_{i+1} - 1$ being equal to zero) while two different ones would give a contribution equal to K, the total contribution of the second term is KU, where U is the number of "kinks" in the chain, i.e. links between two neighboring spins of opposite signs.

For a chain of N spins, the number of microstates corresponding to an energy E can be written as

$$\Omega(N_+, N_-, U) \simeq \binom{N_+}{U/2}\binom{N_-}{U/2}. \tag{378}$$

The formula is derived by taking into account that we have to distribute N_+ spins among $U/2$ groups and N_- among the remaining $U/2$. Each of these distributions gives a binomial term, and, since they are independent, the total number of states is the product of the two binomials. The expression is not exact because we are on a ring, but corrections are however of order N and do not affect the entropy. A slight correction to formula (378) is present for small N_+, N_- and $U/2$, and all these numbers should be indeed reduced by unity.

Introducing $m = M/N$, $u = U/N$ and $\varepsilon = E/N = -m^2/2 + Ku$, one thus finally gets the entropy

$$\begin{aligned}\tilde{s}(\varepsilon, m) &= \frac{1}{N}\ln\Omega = \frac{1}{2}(1+m)\ln(1+m) + \frac{1}{2}(1-m)\ln(1-m) - u\ln u \\ &\quad - \frac{1}{2}(1+m-u)\ln(1+m-u) - \frac{1}{2}(1-m-u)\ln(1-m-u),\end{aligned} \tag{379}$$

which is shown in Fig. 43 for different values of the energy ε and for two values of the nearest-neighbor coupling K.

In the large N-limit, the last step is to maximize the entropy $\tilde{s}(\varepsilon, m)$ with respect to the magnetization m, leading to the final entropy $s(\varepsilon) = \tilde{s}(\varepsilon, m^*)$, where m^* is the equilibrium value. As anticipated, one gets either a second or a first order phase transition depending on the value of the coupling constant K. As usual, and analogously to what has been done in the canonical ensemble, the expansion of $\tilde{s}(\varepsilon, m)$ in power of m is the appropriate procedure to define the critical lines and points. One gets here

$$\tilde{s}(\varepsilon, m) = s_0(\varepsilon) + A_{mc}\, m^2 + B_{mc}\, m^4 + \mathcal{O}(m^6), \tag{380}$$

with the paramagnetic zero magnetization entropy

$$s_0(\varepsilon) = -\frac{\varepsilon}{K}\ln\frac{\varepsilon}{K} - \left(1 - \frac{\varepsilon}{K}\right)\ln\left(1 - \frac{\varepsilon}{K}\right) \tag{381}$$

and the expansion coefficients

$$A_{mc} = \frac{1}{2}\left[\frac{1}{K}\ln\frac{K-\varepsilon}{\varepsilon} - \frac{\varepsilon}{K-\varepsilon}\right] \tag{382}$$

$$B_{mc} = \frac{\varepsilon^3}{12(\varepsilon - K)^3} - \frac{K^2+K}{4(\varepsilon-K)^2} + \frac{1}{8K\varepsilon}. \tag{383}$$

Using these expression, it is straightforward to find the second order phase transition line by requiring that $A_{mc} = 0$ ($B_{mc} < 0$), finding $\beta_c = \exp(-\beta_c K)$, which is the same equation found for the canonical ensemble. Again, as far as second order phase transitions are concerned, the two ensembles are equivalent. The tricritical point is obtained by the condition $A_{mc} = B_{mc} = 0$, which gives $K_{MTP} \simeq -0.359$ and $\beta_{MTP} \simeq 2.21$, which is close but definitely different from $K_{CTP} \simeq -0.317$

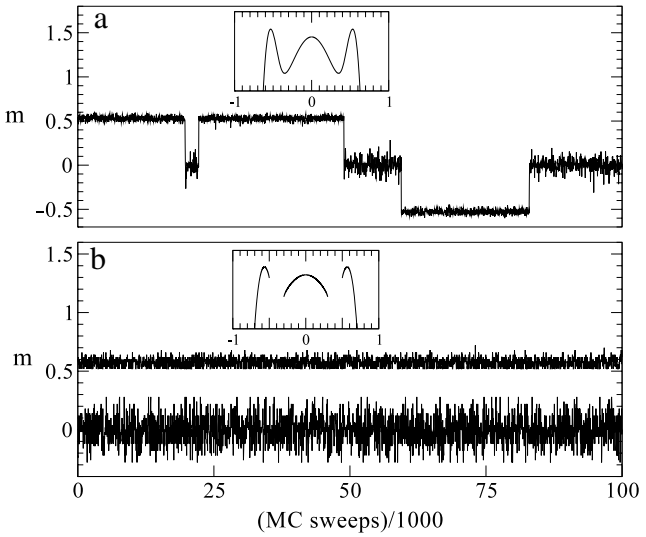


Fig. 44. Time evolution of the magnetization for $K = -0.4$ (a) in the ergodic region ($\varepsilon = -0.318$) and (b) in the non-ergodic region ($\varepsilon = -0.325$). Two different initial conditions are plotted simultaneously. The corresponding entropy curves are shown in the insets.

and $\beta_{CTP} = \sqrt{3}$. The microcanonical first order phase transition line is obtained numerically by equating the entropies of the ferromagnetic and paramagnetic phases. At a given transition energy, there are two temperatures, thus giving a temperature jump as for the BEG and the generalized HMF model. Similarly to mean-field models, also this model exhibits a region of negative specific heat when the phase transition is first order in the canonical ensemble.

6.1.1.4. Equilibrium dynamics: Breaking of ergodicity. Model (369) exhibits breaking of ergodicity, as shown in Ref. [161]. In order to reveal the dynamical consequences of this effect, one has to define a *microcanonical dynamics*. An appropriate one is given by the Creutz algorithm [162], which probes the microstates of the system with energy lower or equal to the energy E. Indeed, there are two definitions of the microcanonical ensemble, that become equivalent in the thermodynamic limit. In the first one, only states contained on an energy shell $(E, E + \mathrm{d}E)$ are counted (this is the definition we used throughout this review); in the second one all the phase-space volume contained within the energy E hypersurface is considered [10]. Creutz algorithm is based on this second definition.

The algorithm is implemented by adding an auxiliary variable, called a *demon*, which has the following properties. One initiates the procedure with the demon at zero energy, while the system has an energy E. One then attempts to flip a spin. The move is accepted if it corresponds to an energy decrease and the excess energy is given to the demon. One then attempts to flip another spin. If the flip decreases the energy, it is accepted, while if it increases the total energy, it is accepted only if the needed energy can be withdrawn from the demon energy. Demon serves really as a bank with deposit and withdrawn: however, the total energy of the "bank" is always non-negative (clearly different from typical banks in the 21st century!).

Creutz dynamics can be used to test the predictions on the breaking of ergodicity discussed in Section 4.5.4. Indeed, as it is already apparent in Fig. 43, not all magnetization values are accessible in certain regions of the (K, ε) plane, which manifests itself in the gaps of the entropy curves shown in Fig. 43. This property is in turn a consequence of the non-convexity of the space of thermodynamic parameters (see Section 2.3)

To demonstrate ergodicity breaking, we show in Fig. 44 the time evolution of the magnetization in two cases: in the first case, where the whole magnetization interval is accessible, one clearly see switches between the metastable state $m^* = 0$ and two stable symmetric magnetized states. On the contrary, in the second example, the metastable state belongs to an interval which is disconnected from the stable one (see the inset for the corresponding entropy curve). The dynamics maintains the system either in the stable or in the metastable interval, depending only on the initial condition. The system is unable to jump to the other interval even when the latter is more stable.

Similar results had been obtained using Hamiltonian dynamics for the generalized HMF model, see Fig. 29.

6.1.2. XY model

6.1.2.1. Introduction. In this subsection we study a generalization of the HMF model, in which the Hamiltonian, besides the mean-field interaction, has a nearest neighbor interaction between rotators. The presence of such a term requires that we specify the properties of the lattice where the model is defined. As in Section 6.1.1, we study a one-dimensional lattice with a ring geometry; the first property allows the use of the transfer integral technique, while the second property, equivalent to the introduction of periodic boundary conditions, is convenient for the calculations, but it is irrelevant in the thermodynamic limit.

The model has been introduced in Ref. [270], and its Hamiltonian is given by

$$H_N = \sum_{i=1}^{N} \frac{p_i^2}{2} + \frac{1}{2N} \sum_{i,j=1}^{N} \left[1 - \cos\left(\theta_i - \theta_j\right)\right] - K \sum_{i=1}^{N} \cos\left(\theta_{i+1} - \theta_i\right), \tag{384}$$

with $\theta_{N+1} = \theta_1$. The parameter K is the coupling constant of the nearest-neighbor interaction. For $K = 0$, the Hamiltonian reduces to the HMF model, that has a second order phase transition at $T_c = 0.5$. The similarities of this model with the Ising model studied in the previous subsection are clear: there is a mean-field interaction and a nearest neighbor interaction. However, there is also a kinetic energy term, which allows one to define a Hamiltonian dynamics. Also in this case interesting properties are found for $K < 0$. Let us first locate the first order transition at $T = 0$. This is done, as for the Ising case, comparing the energy per particle of the fully magnetized $m = 1$ state with that of the staggered non-magnetic $m = 0$ state. The energy density of the former is $\varepsilon = -K$, while that of the latter is $\varepsilon = 1/2 + K$. The magnetized state is therefore favored for $K \geq -1/4$. Qualitatively, it is the same phase diagram as the one observed in Fig. 40.

6.1.2.2. Solutions in the canonical and microcanonical ensembles. We adopt the min–max procedure, described in Section 4.4.4. We recall that, in this procedure, once an expression for the canonical partition function has been obtained in the form of Eq. (155), not only can we compute the canonical free energy $\phi(\beta)$ by the minimization defined in Eq. (156), but we can also obtain the microcanonical entropy $s(\varepsilon)$ by Eq. (160).

Let us first derive the canonical partition function. As for the HMF model, we use the Hubbard–Stratonovich transformation. The model shares with the HMF model the invariance under global rotations, therefore the spontaneous magnetization is defined only in modulus, while there is degeneracy with respect to its direction. We can exploit this fact to simplify slightly the computation; namely, we assume from the beginning that the saddle point value of the variable x_2 in Eq. (130) is 0. The spontaneous magnetization is in the direction of the x_1-axis, but this will not cause any loss of generality. We therefore obtain

$$\begin{aligned} Z(\beta, N) = {} & \frac{N\beta}{2\pi} \exp\left(-\frac{N\beta}{2}\right) \left(\frac{2\pi}{\beta}\right)^{N/2} \\ & \times \int \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_N \mathrm{d}x \exp\left[-\frac{N\beta x^2}{2} + \beta x \sum_{i=1}^{N} \cos\theta_i + \beta K \sum_{i=1}^{N} \cos\left(\theta_{i+1} - \theta_i\right)\right]. \end{aligned} \tag{385}$$

The integral over the θ's is performed applying the transfer integral method. For our one-dimensional geometry this is given by

$$\int \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_N \exp\left[\beta x \sum_{i=1}^{N} \cos\theta_i + \beta K \sum_{i=1}^{N} \cos\left(\theta_{i+1} - \theta_i\right)\right] = \sum_j \lambda_j^N(\beta x, \beta K), \tag{386}$$

where $\lambda_j(\gamma, \sigma)$ is the j-th eigenvalue of the symmetric integral operator

$$(\mathcal{T}\psi)(\theta) = \int \mathrm{d}\alpha \exp\left[\frac{1}{2}\gamma(\cos\theta + \cos\alpha) + \sigma\cos(\theta - \alpha)\right] \psi(\alpha). \tag{387}$$

In the thermodynamic limit, only the largest eigenvalue $\lambda_{\max}$ will contribute to the partition function. This eigenvalue can be computed numerically by a suitable discretization of the integral operator (387). The rescaled free energy is

$$\phi(\beta) = \beta f(\beta) = \inf_x \tilde{\phi}(\beta, x), \tag{388}$$

with

$$\tilde{\phi}(\beta, x) = -\frac{1}{2}\ln\frac{2\pi}{\beta} + \frac{\beta}{2}(1 + x^2) - \lambda_{\max}(\beta x, \beta K), \tag{389}$$

where we have not explicitly written the K dependence of $\tilde{\phi}(\beta, x)$ and $\phi(\beta)$.

Accordingly to the min–max procedure, the microcanonical entropy $s(\varepsilon)$ is given by

$$s(\varepsilon) = \sup_x \tilde{s}(\varepsilon, x), \tag{390}$$

with

$$\tilde{s}(\varepsilon, x) = \inf_\beta \left[\beta\varepsilon - \tilde{\phi}(\beta, x)\right] = \inf_\beta \left[\beta\varepsilon + \frac{1}{2}\ln\frac{2\pi}{\beta} - \frac{\beta}{2}(1 + x^2) + \ln\lambda_{\max}(\beta x, \beta K)\right]. \tag{391}$$

Again the dependence of $\tilde{s}(\varepsilon, x)$ and $s(\varepsilon)$ on K has not been explicitly written. The expressions (389) and (391) are used to study canonical and microcanonical thermodynamic phase diagrams, respectively, and then to check if there is ensemble

inequivalence. As for the Ising case, to obtain the critical lines on the (T, K) plane, one has to expand $\tilde{\phi}(\beta, x)$ and $\tilde{s}(\varepsilon, x)$ in powers of x. The power expansion of $\lambda_{\max}(\beta x, \beta K)$ can be obtained explicitly as a function of modified Bessel functions. We will avoid here the details of the calculation leading to the final result.

Taking into account that our system is invariant under the symmetry $\theta \to -\theta$, the expansion of $\tilde{\phi}(\beta, x)$ will have only even powers. Let us write explicitly the K dependence of the expansion coefficients

$$\tilde{\phi}(\beta, x) = \phi_0(\beta, K) + \phi_1(\beta, K)\, x^2 + \phi_2(\beta, K)\, x^4 + o(x^6). \tag{392}$$

The canonical second order transition line in the (T, K) plane is determined by

$$\phi_1(\beta, K) = 0, \quad \text{with } \phi_2(\beta, K) > 0. \tag{393}$$

Inserting this expansion in the first equality in Eq. (391) and minimizing with respect to β, one obtains

$$\varepsilon = \phi_0'(\beta, K) + \phi_1'(\beta, K)\, x^2 + \phi_2'(\beta, K)\, x^4 + o(x^6), \tag{394}$$

where the prime denotes derivation with respect to β. This equation gives an expansion of the form

$$\beta(\varepsilon, K, x) = \beta_0(\varepsilon, K) + \beta_1(\varepsilon, K)\, x^2 + \beta_2(\varepsilon, K)\, x^4 + o(x^6). \tag{395}$$

Using this expression, we obtain the expansion of the form

$$s(\varepsilon, m) = s_0(\varepsilon, K) + s_1(\varepsilon, K) m^2 + s_2(\varepsilon, K) m^4 + o(m^6), \tag{396}$$

where

$$s_0(\varepsilon, K) = \beta_0(\varepsilon, K)\varepsilon - \phi_0(\beta_0(\varepsilon, K)) \tag{397}$$

$$s_1(\varepsilon, K) = \beta_1(\varepsilon, K)\varepsilon - \phi_0'(\beta_0(\varepsilon, K))\beta_1(\varepsilon, K) - \phi_1(\beta_0(\varepsilon, K)) \tag{398}$$

$$\begin{aligned} s_2(\varepsilon, K) = {} & \beta_2(\varepsilon, K)\varepsilon - \frac{1}{2}\phi_0''(\beta_0(\varepsilon, K))\beta_1^2(\varepsilon, K) - \phi_0'(\beta_0(\varepsilon, K))\beta_2(\varepsilon, K) \\ & - \phi_1'(\beta_0(\varepsilon, K))\beta_1(\varepsilon, K) - \phi_2(\beta_0(\varepsilon, K)). \end{aligned} \tag{399}$$

The microcanonical second order transition line is determined by

$$s_1(\varepsilon, K) = 0, \quad \text{with } s_2(\beta, K) < 0. \tag{400}$$

Now we can use the following equalities, that are obtained by inserting back (395) into (394)

$$\phi_0'(\beta_0(\varepsilon, K)) = \varepsilon \tag{401}$$

$$\phi_1'(\beta_0(\varepsilon, K)) = -\phi_0''(\beta_0(\varepsilon, K))\beta_1(\varepsilon, K). \tag{402}$$

Eq. (401) implies that $\beta_0(\varepsilon, K)$ is the temperature on the second order transition line. One obtains

$$s(\varepsilon, x) = \beta_0(\varepsilon, K)\varepsilon - \phi_0(\beta_0(\varepsilon, K)) - \phi_1(\beta_0(\varepsilon, K))x^2 - \left[-\frac{1}{2}\phi_0''(\beta_0(\varepsilon, K))\beta_1^2(\varepsilon, K) + \phi_2(\beta_0(\varepsilon, K))\right]x^4 + o(x^6). \tag{403}$$

Therefore, the microcanonical second order transition line is determined by

$$\phi_1(\beta_0(\varepsilon, K)) = 0, \quad \text{with } \frac{1}{2}\phi_0''(\beta_0(\varepsilon, K))\beta_1^2(\varepsilon, K) - \phi_2(\beta_0(\varepsilon, K)) < 0. \tag{404}$$

The concavity of $\phi(\beta)$ implies that, again on the critical line, $\phi_0''(\beta_0(\varepsilon, K)) < 0$ (this is valid in models where parity with respect to x implies that $\tilde{\phi}_{\beta x} = 0$ for $m = 0$). Comparing Eqs. (393) and (404), and taking into account the last observation, we see that, as we expected from the general discussion about ensemble equivalence at the end of Section 4.1.3, the points belonging to the canonical second order transition line also belong to the microcanonical one. The microcanonical line, however, goes beyond and includes more points, as we have already seen in several models, since by continuity the inequality in (404) is also satisfied for some points where $\phi_2(\beta_0(\varepsilon, K)) < 0$.

The study of the canonical and microcanonical critical lines involves the analytic determination of $\lambda_{\max}$, which, in the limit $x \to 0$, can be performed using perturbation theory for Hermitian operators [271]. For both ensembles, one gets

$$\frac{I_1(\beta K)}{I_0(\beta K)} = \frac{4 - \beta}{2\beta + 4}, \tag{405}$$

where I_0 and I_1 are modified Bessel functions. The canonical tricritical point turns out to be $K_{CTP} \simeq -1.705$ and $T_{CTP} \simeq 0.267$, while microcanonical tricritical point is at $K_{MTP} \simeq -0.182$ and $T_{MTP} \simeq 0.234$. In Fig. 45, we plot the relevant part of the phase diagram on the (K, T) plane. As commented above, it is shown how the microcanonical critical line (ending at the microcanonical tricritical point MTP) extends beyond the canonical critical line, which ends at the canonical tricritical point CTP.

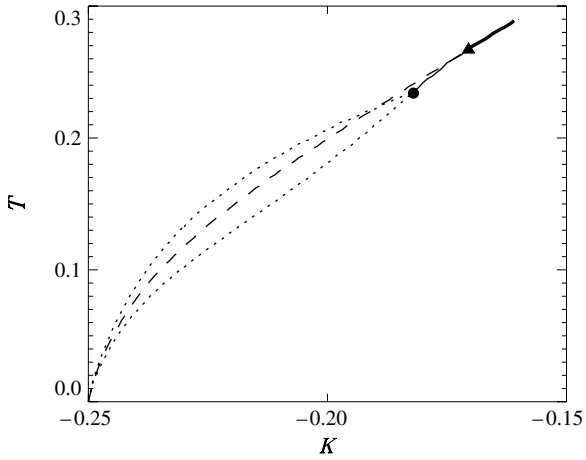


Fig. 45. Canonical and microcanonical (K, T) phase diagram of the one-dimensional XY model with mean-field and nearest-neighbor interactions. The canonical critical line (bold solid line) ends at the tricritical point CTP indicated by a triangle; then the transition becomes first order (dashed line). The microcanonical second order transition line coincides with the canonical one at large K (bold solid line). It continues at lower K (light solid line) down to the tricritical point MTP, indicated by a filled circle; then the transition becomes first order, with a branching of the transition line (dotted lines), giving the two extremes of the temperature jump.

6.1.2.3. Ergodicity breaking. We here prove that model (384) shows ergodicity breaking, similarly to the Ising long- plus short-range model. Let us first study separately the bounds of the short-range term

$$\frac{1}{N}\sum_{i=1}^{N}\cos\left(\theta_{i+1}-\theta_i\right). \tag{406}$$

The study of maxima and minima of this term can be reduced to those configurations where the spins are aligned parallel or antiparallel with a given direction. There is no loss of generality, due to rotational symmetry, to then choose this direction as the x-axis. Therefore, only the x component of m is non-zero and in the following we will restrict to $m = m_x \geq 0$. It turns out that, in the thermodynamic limit, both extrema are attained with a fraction $(1+m)/2$ of rotators parallel with the x-axis and a fraction $(1-m)/2$ antiparallel with the x-axis, for all given values of m. The maximum is 1 and is achieved when the parallel rotators and the antiparallel rotators are grouped in two separated blocks. The minimum $2m-1$, is instead attained when the antiparallel rotators are all isolated, which is possible since we are considering $m \geq 0$.

Now we have to distinguish the two cases $K > 0$ and $K < 0$. In the positive case, the minimum of this contribution to the energy density is $-K$ (thus actually independent of m), while in the negative case it is $-K(2m-1)$. We obtain the minimum of the total potential energy per particle just by adding $(1-m^2)/2$. This is actually the minimum of the total energy per particle ε, since the kinetic energy is positive definite. We finally obtain that the minimum of the energy per particle is

$$\varepsilon_{\min}^{K>0}(m) = \frac{1-m^2}{2} - K \tag{407}$$

for $K > 0$, and

$$\varepsilon_{\min}^{K<0}(m) = \frac{1}{2}(1-m^2) - K(2m-1) \tag{408}$$

for $K < 0$. Plotting the two functions $\varepsilon_{\min}^{K>0}$ and $\varepsilon_{\min}^{K<0}$ we will find the accessible m values for a given energy ε. The functions are plotted in Fig. 46 for two representative values: $K = 0.2$ and $K = -0.2$. For positive K, we see that below the energy $-K + 1/2$ the system cannot have all magnetizations, and values of m near 0 are inaccessible. We note however that the region of accessible values of m, for any energy, is connected; therefore there is no ergodicity breaking. Nevertheless, we remark that the accessible region in the (ε, m) plane is not convex, as it can happen only for long-range interactions. For negative K, we note instead, that for a given energy interval ($K + 1/2 < \varepsilon < 2K^2 + K + 1/2$ for $K > -1/4$ and $-K < \varepsilon < 2K^2 + K + 1/2$ for $K < -1/4$), the attainable values of m are separated in two disconnected regions. Therefore for $K < 0$ we have a breaking of ergodicity.

6.2. Weakly decaying interactions

All the models that have been considered in previous sections have an infinite range term in the Hamiltonian. This undoubtedly limits the applicability of our analysis to realistic physical systems, where interactions decay with the distance. We present below four models of this kind which are direct generalizations of models presented above and represent

a step forward in the study of long-range weakly decaying interactions. Some of the models are exactly solvable and reproduce features of the phenomenology observed for infinite range systems (ensemble inequivalence, negative specific heat, ergodicity breaking, etc.).

6.2.1. α-Ising model

Let us consider the one-dimensional α-Ising Hamiltonian

$$H_N = \frac{J}{N^{1-\alpha}} \sum_{i>j=1}^{N} \frac{1 - S_i S_j}{|i-j|^\alpha}, \tag{409}$$

where $J > 0$ and spins $S_i = \pm 1$ sit on a one-dimensional lattice with either free or periodic boundary conditions (in the latter case, $|i-j|$ is the minimal distance along the ring). The $N^{\alpha-1}$ prefactor is introduced in order to have an extensive energy as explained in Section 2.1. This model has been first introduced by Dyson [272] and studied for the "integrable" case, $\alpha > 1$, in the canonical ensemble without the $N^{\alpha-1}$ prefactor. We show here that it is possible to obtain an exact microcanonical solution using large deviation theory when $0 \leq \alpha < 1$ [114,127,273]. Moreover, the study of this model gives also the opportunity to emphasize the important role played by boundary conditions when the interactions are long-range.

We adopt the same scheme described in Section 4.5 to obtain the solution of model (409). The method can be generalized to lattices of higher dimension.

In the first step, the Hamiltonian H_N is rewritten in terms of global variables by introducing a *coarse-graining*. Let us divide the lattice in K boxes, each with $n = N/K$ sites, and let us introduce the average magnetization in each box m_k, $k = 1 \dots K$. In the limit $N \to \infty, K \to \infty, K/N \to 0$, the magnetization becomes a continuous function $m(x)$ of the $[0, 1]$ interval. After a long but straightforward calculation, described in Ref. [127], it is possible to express H_N as a functional of $m(x)$

$$H_N = NH[m(x)] + o(N), \tag{410}$$

where

$$H[m(x)] = \frac{J}{2} \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \frac{1 - m(x)m(y)}{|x-y|^\alpha}. \tag{411}$$

The estimation is uniform on all configurations.

In the second step, we evaluate the probability to get a given magnetization m_k in the k-th box from all *a priori* equiprobable microscopic configurations. This probability obeys a local large deviation principle $P(m_k) \propto \exp[n\tilde{s}(m_k)]$, with

$$\tilde{s}(m_k) = -\frac{1+m_k}{2} \ln \frac{1+m_k}{2} - \frac{1-m_k}{2} \ln \frac{1-m_k}{2}. \tag{412}$$

Since the microscopic random variables in the different boxes are independent and no global constraint has been imposed, the probability of the full global variable $(m_1, \dots, m_K)$ can be expressed in a factorized form as

$$P(m_1, m_2, \dots, m_K) = \prod_{i=1}^{K} P(m_i) \simeq \prod_{i=1}^{K} \mathrm{e}^{n\tilde{s}(m_i)} \tag{413}$$

$$= \exp\left[nK \sum_{i=1}^{K} \frac{\tilde{s}(m_i)}{K} \right] \tag{414}$$

$$\simeq \mathrm{e}^{N\bar{s}[m(x)]}, \tag{415}$$

where $\bar{s}[m(x)] = \int_0^1 \tilde{s}(m(x))\mathrm{d}x$ is the entropy functional associated to the global variable $m(x)$. Large deviation techniques rigorously justify these calculations [274], proving that entropy is proportional to N, also in the presence of long-range interactions. This result is independent of the specific model considered.

In the third step, we formulate the variational problem in the microcanonical ensemble to get the entropy

$$s(\varepsilon) = \sup_{m(x)} \left(\bar{s}[m(x)] \,|\, \varepsilon = H[m(x)] \right). \tag{416}$$

Let us remark that this optimization problem has to be solved in a functional space. In general, this has to be done numerically, taking into account boundary conditions. In the case of free boundary conditions, the only available solutions are numerical. An example of a maximal entropy magnetization profile obtained for free boundary conditions is shown in Fig. 47 for different values of α. The profile becomes more inhomogeneous when increasing α (for $\alpha = 0$ one recovers the mean-field result with a homogeneous profile).

In the following, we will treat the periodic boundary conditions case, for which analytical results can be obtained. Both entropy and free energy can be obtained in analytical form for homogeneous magnetization profiles, which have been shown to be locally stable in both the high-temperature and the low-temperature phase [127]. It has indeed been proven that for

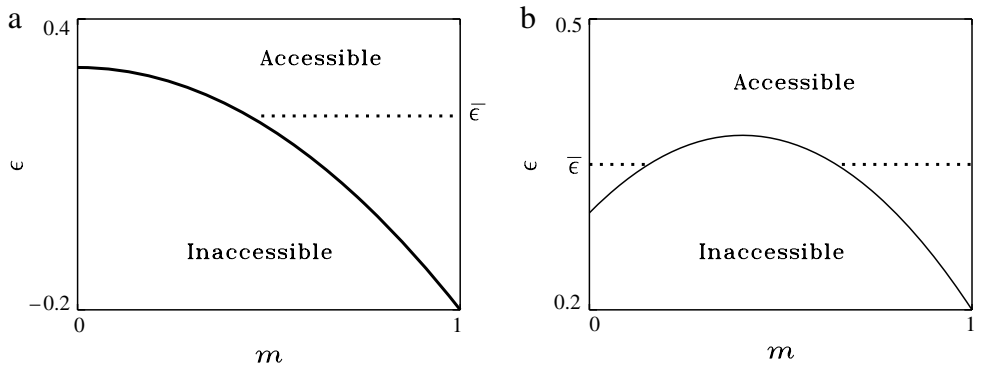


Fig. 46. Accessible regions in the (ε, m) plane for the one-dimensional XY model with both mean-field and nearest neighbor interactions. Solid curves represent (a) $\varepsilon_{\min}^{K>0}(m)$ (407) and (b) $\varepsilon_{\min}^{K<0}(m)$ (408). The accessible and inaccessible regions in parameter space are indicated. (a) $K = 0.2$: no breaking of ergodicity is present since at fixed energy $\bar{\varepsilon}$ the accessible interval of m is connected (dotted line); (b) $K = -0.2$: there is breaking of ergodicity for values of ε between 0.30 and 0.38. For instance for $\bar{\varepsilon} = 0.35$ there are two disconnected intervals of accessible magnetizations (dotted lines).

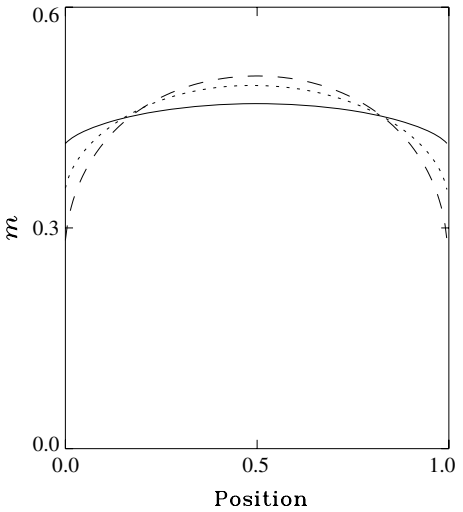


Fig. 47. Equilibrium magnetization profile for the α-Ising model with free boundary conditions at an energy density $\varepsilon = 0.1$ for $\alpha = 0.2$ (solid line), $\alpha = 0.5$ (dotted line) and $\alpha = 0.8$ (dashed line).

$\beta < \beta_c = (1-\alpha)/(J2^{\alpha})$ there is a unique global maximum of $s(\varepsilon)$, corresponding to a constant zero magnetization profile. The variational problem (416), where $\bar{s}$ is defined in Eqs. (412), (414) and (415), leads to the consistency equation

$$\tanh^{-1}(m(x)) = \beta J \int_0^1 \frac{m(y)}{|x-y|^{\alpha}} \mathrm{d}y, \tag{417}$$

where β is a Lagrange multiplier. For $\beta > \beta_c$, we restrict ourselves to constant magnetization profiles, which are locally stable, i.e. close non-constant profiles have a smaller entropy. In this case, using the relations $\int_0^1 \mathrm{d}x|x-y|^{-\alpha} = 2^{\alpha}/(1-\alpha)$ and $\varepsilon_{\max} = 1/(2\beta_c)$ one can obtain the magnetization vs. energy curve $m = \pm\sqrt{1-\varepsilon/\varepsilon_{\max}}$, see Fig. 48(a). Moreover, fixing the energy implies fixing the magnetization and, consequently, the Lagrange multiplier β in Eq. (417). Expressing the magnetization in terms of the energy in the entropy formula (416) allows us to derive the caloric curve, see Fig. 48(b) (solid line). The limit temperature β_c (dotted line) is attained at zero magnetization, which is a boundary point.

In the canonical ensemble, one has to solve the variational problem (195). This leads to exactly the same consistency equation (417), where the Lagrange multiplier is replaced by the inverse temperature β. Solving this consistency equation for $\beta > 0$, one finds a zero magnetization for $\beta < \beta_c$ and a non-vanishing one for $\beta > \beta_c$. One can also derive the canonical caloric curve, which is reported in Fig. 48(b) and superposes to the microcanonical caloric curve from $\beta = \infty$ down to β_c while it is represented by the dashed line for $\beta < \beta_c$. It follows that in the region $[0, \beta_c]$, the two ensembles are not equivalent. In this case, a single microcanonical state at $\varepsilon_{\max}$ corresponds to many canonical states with canonical inverse temperatures in the range $[0, \beta_c[$. Thus, in such a case, the canonical inverse temperature is not equal to the microcanonical one. In the microcanonical ensemble, the full high temperature region is absent and, therefore, no phase transition is present or, in other terms, the phase transition is at the boundary of the accessible energy values. The entropy is always concave,

hence no inequivalence can be present in the allowed energy range, apart from the boundaries. This situation is often called partial equivalence [99,152,275]. Partial equivalence persists for all α values below one, and is removed only for $\alpha = 1$ when $\varepsilon_{\max} \to \infty$ and $\beta_c \to 0$: the phase transition is not present in both ensembles and the system is always in its magnetized phase. The main drawback of this analysis is the difficulty of obtaining analytical solutions of Eq. (417) for non-constant magnetization profiles, which is the typical situation when boundary conditions are not periodic.

6.2.2. α-HMF model

Let us consider now a generalization of the HMF model (126) which has been originally proposed in Ref. [276,277]. The interaction between two rotators has the same form as in the original model, but the coupling constant is a weakly decaying function of the distance between the two lattice sites where the rotators sit. Note that we consider the lattice in general dimension d. In Ref. [144], the generalized case of n-vector spin models has been analyzed, including the HMF model (for $n = 2$) and the Ising spins (for $n = 1$). Only periodic boundary conditions have been considered, as we will do here.

The Hamiltonian of the α-HMF model is

$$H_N = \sum_{i=1}^{N} \frac{p_i^2}{2} + \frac{1}{2\widetilde{N}} \sum_{i,j=1}^{N} \frac{1 - \cos(\theta_i - \theta_j)}{r_{ij}^{\alpha}}, \tag{418}$$

where r_{ij} is the minimal distance between lattice sites i and j (minimal distance convention in periodic boundary conditions). We consider here cases for which the exponent α, determining the decay with distance of the coupling constant between rotators, is less than the spatial dimension d, so that the system is long-range as shown in Section 2.2. The $\widetilde{N}$ prefactor is introduced in order to have an extensive energy as explained in Section 2.1. We show below that, in any spatial dimension d and for each $\alpha < d$, the thermodynamics of this model is the same as that of the HMF model; in particular, there is a second order phase transition at the inverse temperature $\beta = 2$.

Let us introduce the matrix of the coupling constants $J_{ij} = 1/r_{ij}^{\alpha}$. Since the diagonal elements of this matrix are not defined, we assign to them the common arbitrary value b. The prefactor is

$$\widetilde{N} = \sum_{j=1}^{N} J_{ij} \tag{419}$$

that does not depend on i, because we adopt the minimal distance convention for r_{ij}. It is clear that in the thermodynamic limit $N \to \infty$, the prefactor $\widetilde{N}$ behaves as $N^{1-\alpha/d}$. The arbitrary parameter b, which appears only in $\widetilde{N}$, becomes negligible in thermodynamic limit. Its introduction is however useful for the computation, as it will become clear.

The most relevant steps of the solution in the canonical ensemble (we refer the reader to Ref. [144] for details) can be obtained starting from the partition function

$$Z(\beta, N) = \left(\frac{2\pi}{\beta}\right)^{N/2} \int \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_N \exp\left(-\frac{\beta}{2\widetilde{N}} \sum_{i,j=1}^{N} J_{ij}\left[1 - \cos(\theta_i - \theta_j)\right]\right) \tag{420}$$

$$= \left(\frac{2\pi}{\beta}\right)^{N/2} \exp\left(-\frac{N\beta}{2}\right) \int \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_N \exp\left(\sum_{i,j=1}^{N}\left[\cos\theta_i R_{ij} \cos\theta_j + \sin\theta_i R_{ij} \sin\theta_j\right]\right), \tag{421}$$

where we have introduced the matrix $R_{ij} = \beta J_{ij}/(2\widetilde{N})$ and used the property $\sum_{ij} J_{ij} = N\widetilde{N}$. In order to use the Hubbard–Stratonovich transformation, one diagonalizes the expression in the second exponential. Denoting by R the symmetric matrix with elements R_{ij}, and by V the unitary matrix that reduces R to its diagonal form $D = VRV^{\mathrm{T}}$, one gets

$$\sum_{i,j=1}^{N} \cos\theta_i R_{ij} \cos\theta_j = \sum_{i=1}^{N} R_i a_i^2, \tag{422}$$

where the quantities R_i are the real eigenvalues of the matrix R, and $a_i = \sum_{j=1}^{N} V_{ij} \cos\theta_j$. An analogous expression holds for the sum with the sines in Eq. (421). In order to express $\exp(R_i a_i^2)$ using the Hubbard–Stratonovich transformation, the eigenvalues R_i must be positive. By choosing a sufficiently large value of b, but in any case of order 1, all eigenvalues R_i are positive. One can therefore write

$$\exp\left(\sum_{i=1}^{N} R_i a_i^2\right) = \frac{1}{\sqrt{(4\pi)^N \det R}} \int \mathrm{d}y_1 \ldots \mathrm{d}y_N \exp\left[\sum_{i=1}^{N}\left(-\frac{y_i^2}{4R_i} + a_i y_i\right)\right] \tag{423}$$

$$= \frac{1}{\sqrt{(4\pi)^N \det R}} \int \mathrm{d}z_1 \ldots \mathrm{d}z_N \exp\left[-\frac{1}{4}\sum_{i,j} z_i \left(R^{-1}\right)_{ij} z_j + \sum_{i=1}^{N} z_i \cos\theta_i\right]. \tag{424}$$

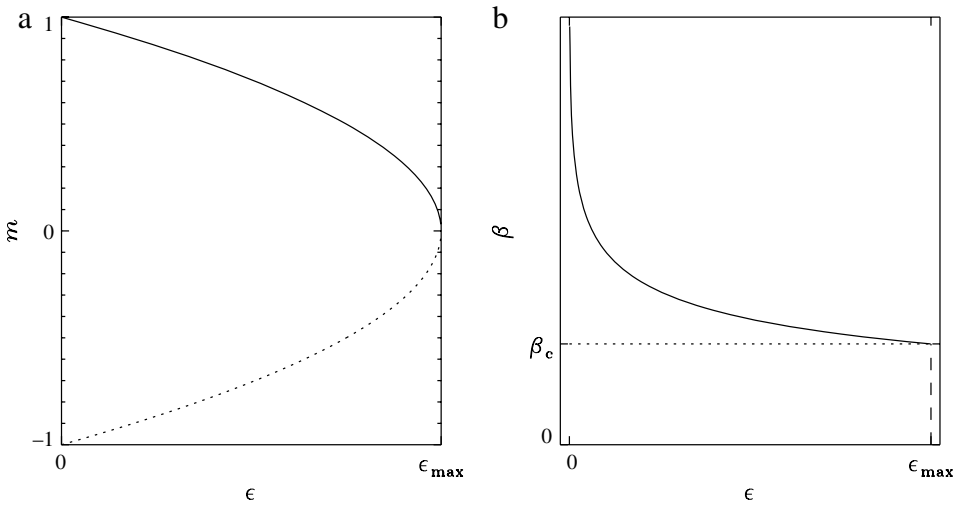


Fig. 48. (a) Equilibrium magnetization in the allowed energy range in the microcanonical ensemble for the α-Ising model with $\alpha = 0.5$; the negative branch is also reported with a dotted line. (b) Inverse temperature versus energy in the microcanonical ensemble (solid line). The canonical ensemble result superposes to the microcanonical one in the interval $[\beta_c, \infty)$ and is represented by a dashed line for $\beta \in [0, \beta_c]$. β_c is then the inverse critical temperature in the canonical ensemble. In the microcanonical ensemble, no phase transition is present.

In this N-dimensional Hubbard–Stratonovich transformation, we have used that $\det R = \prod_{i=1}^{N} R_i$, we have introduced the unitary change of integration variables defined by $y_i = \sum_{j=1}^{N} V_{ij} z_j$ and we have indicated with R^{-1} the inverse matrix of R. Using the analogous expression for the term with the sines, we arrive, after performing the integration over the angles θ, to the following expression for the partition function

$$Z(\beta, N) = \left(\frac{2\pi}{\beta}\right)^{N/2} \frac{\exp(-N\beta/2)}{(4\pi)^N \det R} \int \mathrm{d}z_1 \dots \mathrm{d}z_N \mathrm{d}z'_1 \dots \mathrm{d}z'_N \times \exp\left[-\frac{1}{4}\sum_{i,j} z_i \left(R^{-1}\right)_{ij} z_j - \frac{1}{4}\sum_{i,j} z'_i \left(R^{-1}\right)_{ij} z'_j + \sum_{i=1}^{N} \ln I_0\left(\sqrt{z_i^2 + z_i'^2}\right)\right]. \tag{425}$$

This integral can be computed using the saddle point method, since most eigenvalues R_i vanish in the thermodynamic limit [144]. Introducing $\rho_i = \sqrt{z_i^2 + z_i'^2}$, the stationary points are solutions of the following consistency equations

$$\frac{1}{2}\sum_{j=1}^{N} \left(R^{-1}\right)_{ij} z_j = \frac{I_1(\rho_i)}{I_0(\rho_i)} \frac{z_i}{\rho_i} \tag{426}$$

$$\frac{1}{2}\sum_{j=1}^{N} \left(R^{-1}\right)_{ij} z'_j = \frac{I_1(\rho_i)}{I_0(\rho_i)} \frac{z'_i}{\rho_i}, \tag{427}$$

which, after inversion with respect to z_i, z'_i, become

$$z_i = 2\sum_{j=1}^{N} R_{ij} \frac{I_1(\rho_j)}{I_0(\rho_j)} \frac{z_j}{\rho_j} \tag{428}$$

$$z'_i = 2\sum_{j=1}^{N} R_{ij} \frac{I_1(\rho_j)}{I_0(\rho_j)} \frac{z'_j}{\rho_j}. \tag{429}$$

Let us first look for homogeneous solutions, i.e., for solutions in z_i, z'_i that do not depend on i. The previous system reduces to the following pair of equations for $z \equiv z_i$ and $z' \equiv z'_i$

$$z = \beta \frac{I_1(\rho)}{I_0(\rho)} \frac{z}{\rho} \quad \text{and} \quad z' = \beta \frac{I_1(\rho)}{I_0(\rho)} \frac{z'}{\rho}, \tag{430}$$

where we have used that $2\sum_{j=1}^{N} R_{ij} = \beta$ for each i. By expressing $z = \rho\cos\gamma$ and $z' = \rho\sin\gamma$, we see that these two equations determine only the modulus $\rho \equiv \sqrt{z^2 + z'^2}$, leaving the angle γ undetermined. They thus reduce to a single

equation for the modulus

$$\rho = \beta \frac{I_1(\rho)}{I_0(\rho)}. \tag{431}$$

This equation, after the change of variable $\rho = \beta m$, coincides with Eq. (134) of the HMF model.

In order for this homogeneous stationary point to be the relevant one for the evaluation of the integral in Eq. (425), it is necessary that the exponential has an absolute maximum at this stationary point. One can easily show that the homogeneous stationary point is a local maximum [144]. Here, we will directly prove that it is also a global maximum with respect to any other possible inhomogeneous stationary point.

Let us then go back to Eqs. (428) and (429) and suppose that they have an inhomogeneous solution, considering as a first case a solution with different values for the moduli $\rho_i \equiv \sqrt{z_i^2 + z_i'^2}$. If k is the site with the largest modulus ρ_k, let us choose the x-axis of the XY rotators such that $z_k = \rho_k$ and $z_k' = 0$. Then one can write

$$z_k = \rho_k = 2\sum_{j=1}^{N} R_{kj} \frac{I_1(\rho_j)}{I_0(\rho_j)} \frac{z_j}{\rho_j} \leq 2\sum_{j=1}^{N} R_{kj} \frac{I_1(\rho_j)}{I_0(\rho_j)} \tag{432}$$

$$< 2\sum_{j=1}^{N} R_{kj} \frac{I_1(\rho_k)}{I_0(\rho_k)} = \beta \frac{I_1(\rho_k)}{I_0(\rho_k)}, \tag{433}$$

where use has been made of the monotonicity of the function I_1/I_0. We emphasize that the last inequality is strict since, by hypothesis, not all moduli ρ_i are equal. From the properties of the function I_1/I_0, it follows from the last inequality that also the largest modulus ρ_k of this inhomogeneous stationary point is smaller than that of the homogeneous solution. One gets the same conclusion for a possible stationary point with all equal moduli but with different directions; in that case inequality (433) becomes an equality, but inequality (433) is strict. The proof is completed by showing that the exponent in Eq. (425), if computed at a stationary point, has a positive derivative with the respect the moduli ρ_i of the stationary point. In fact, it can be easily shown that at a stationary point the exponent in Eq. (425), let us call it A, can be written as a function of the moduli ρ_i of the stationary point in the following way

$$A = \sum_{i=1}^{N} \left[-\frac{1}{2}\rho_i \frac{I_1(\rho_i)}{I_0(\rho_i)} + \ln I_0(\rho_i) \right]. \tag{434}$$

Its derivative with respect to ρ_i is

$$\frac{\partial A}{\partial \rho_i} = \frac{1}{2}\left[\frac{I_1(\rho_i)}{I_0(\rho_i)} - \rho_i \frac{\partial}{\partial \rho_i} \frac{I_1(\rho_i)}{I_0(\rho_i)} \right]. \tag{435}$$

The concavity of the function I_1/I_0, for positive values of the argument, assures that the right hand side is positive. This concludes the proof.

We conclude that for each dimension d and for each decaying exponent $\alpha < d$, the free energy of the α-HMF model with periodic boundary conditions is the same as that of the HMF model. Maybe the derivation of the solution in the microcanonical ensemble does not pose particular technical problems [278].

6.2.3. *One-dimensional gravitational models*

A direct study of the full three-dimensional N-body gravitational dynamics is particularly heavy [279] and special purpose machines have been built to this aim [280]. Therefore, lower dimensional models have been introduced to describe gravitational systems with additional symmetries. For instance, the gravitational sheet model, describing the motion of infinite planar mass distributions perpendicularly to their surface, has been considered [122]. Although this model shows interesting behaviors [281,282], the specific heat is always positive and no phase transition is present. Long ago, Hénon [283] introduced a system of concentric spherical mass shells that interact via gravitational forces. This model has been recently shown to display phase transitions and ensemble inequivalence in a mean-field limit [120].

In this section we discuss the phase diagram, in both the microcanonical and the canonical ensemble, of the Self-Gravitating-Ring (SGR) [284] model. In this model, particle motion is constrained on a ring and particles interact via a true 3D Newtonian potential. The Hamiltonian of the SGR model is

$$H_N = \frac{1}{2}\sum_{i=1}^{N} p_i^2 + \frac{1}{2N}\sum_{i,j} V_\delta(\theta_i - \theta_j), \tag{436}$$

$$\text{with } V_\delta(\theta_i - \theta_j) = -\frac{1}{\sqrt{2}} \frac{1}{\sqrt{1 - \cos(\theta_i - \theta_j) + \delta}}, \tag{437}$$

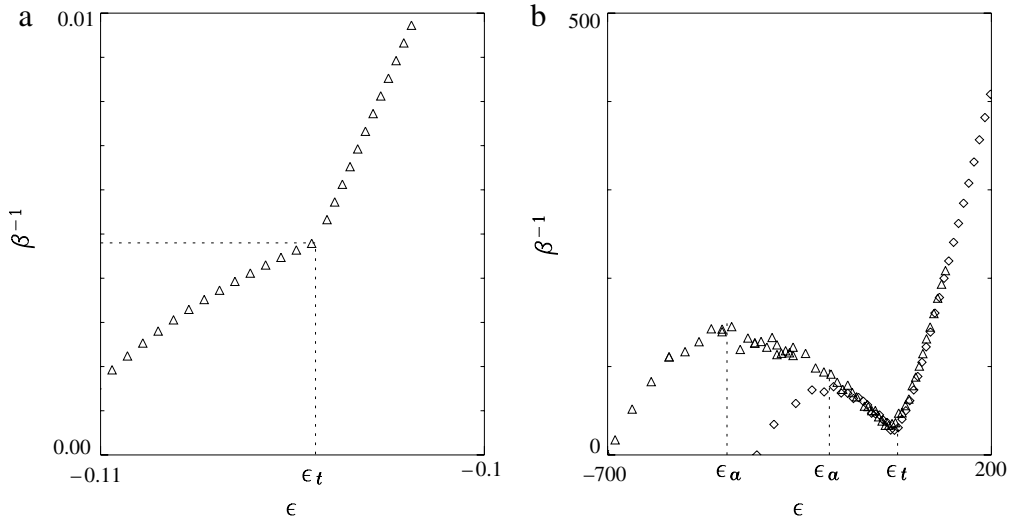


Fig. 49. Caloric curves of the Self Gravitating Ring (SGR) model obtained from numerical simulations of Hamiltonian (436). Panel (a) refers to the softening parameter value $\delta = 10$, for which a second order phase transition appears at ε_t. No backbending of the caloric curve, indicating a negative specific heat, is present. Simulations were performed for $N = 100$. Panel (b) shows the caloric curves for two different values of the softening parameter, $\delta_1 = 1.0 \times 10^{-6}$ and $\delta_2 = 2.5 \times 10^{-7}$, and $N = 100$. The transition is here first order in the microcanonical ensemble. A negative specific heat phase appears for $\varepsilon_a < \varepsilon < \varepsilon_t$, and expands as the softening parameter is reduced.

where δ is the softening parameter, which is introduced, as usual, in order to avoid the divergence of the potential at short distances. Taking the large δ limit, the potential becomes

$$V_\delta = \frac{1}{\sqrt{2\delta}}\left[\frac{1-\cos\left(\theta_i-\theta_j\right)}{2\delta}-1\right]+O(\delta^{-2}), \tag{438}$$

which is that of the Hamiltonian Mean-Field (HMF) model (126). Since the interaction is regularized at short distances, a global entropy maximum always exists, and thermodynamics is well defined in the mean-field limit. In a situation close to that of the HMF model e.g. for $\delta = 10$, the caloric curve determined from microcanonical numerical simulations is reported in Fig. 49(a). In the homogeneous phase $\varepsilon > \varepsilon_t(\delta)$, the caloric curve is almost linear, while in the clustered phase $\varepsilon < \varepsilon_t(\delta)$, it is bent downwards. Nonetheless, temperature always grows with energy and one does not observe any negative specific heat energy range. However, as it happens for 3D Newtonian gravity simulations [279], when one reduces the softening parameter, a negative specific heat phase develops. For instance, in Fig. 49(b), we show two cases at small δ where three phases can be identified [284]:

- A low-energy clustered phase for $\varepsilon < \varepsilon_a$, where ε_a is defined as the energy at which the negative specific heat region begins.
- An intermediate-energy phase, $\varepsilon_a(\delta) < \varepsilon < \varepsilon_t(\delta)$, with negative specific heat.
- A high-energy gaseous phase for $\varepsilon_t(\delta) < \varepsilon$.

The clustered phase is created by the presence of softening, without which the particles would fall into the zero distance singularity. In the gas phase, the particles are hardly affected by the potential and behave as almost free particles. The intermediate phase is expected to show the characters of gravity, persisting and even widening in the $\delta \to 0$ limit.

It is shown analytically in Ref. [285] that ensembles are not equivalent and a phase of negative specific heat and temperature jumps in the microcanonical ensemble appears in a wide intermediate energy region, if the softening parameter is small enough. In Fig. 50, that is drawn for $\delta = 10^{-5}$, we show all the features that we had already commented in the context of the study of the phase transition of the BEG model (Section 4.2). In particular, we would like to draw the attention of the reader to the similarities of Fig. 50(b) with Fig. 13(a).

The phase transition changes from second to first order at a tricritical point, whose location is not the same in the two ensembles. In Fig. 51, we represent the δ-dependence of the critical energy ε_t and the so-called coexistence region. The transition line separating the low energy and high energy phases has a tricritical point: phase transitions are first order at small δ and second order at large δ.

6.2.4. Dipolar interactions in a ferromagnet

In Section 3, where we have discussed several physical examples of long-range interacting systems, we have remarked that dipolar interactions are marginal as far as the long-range character of the interaction is concerned. Indeed, in this case, the exponent governing the decay of the interaction potential is $\alpha = d = 3$. In spite of this, we will show here that magnetic

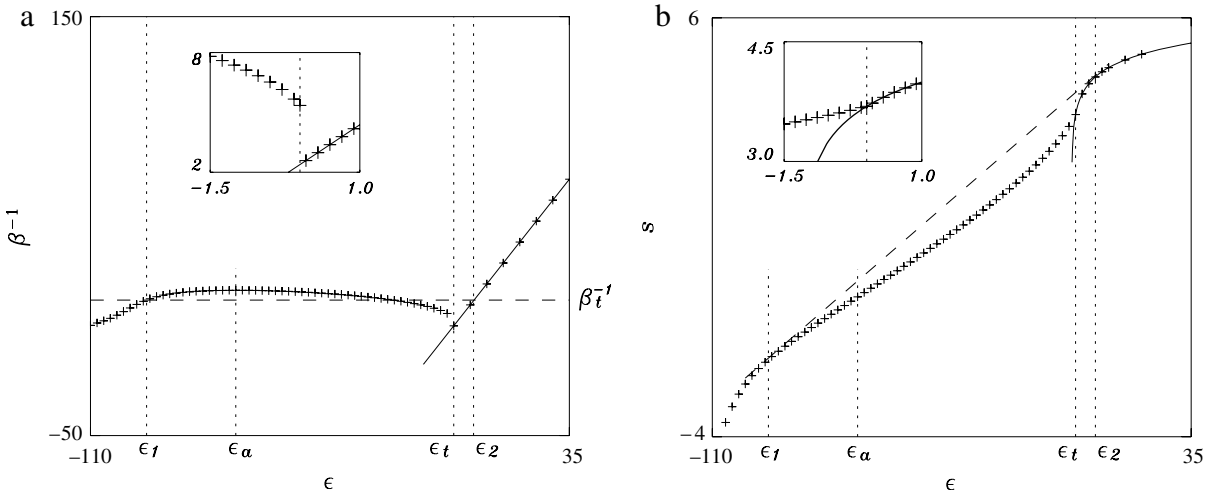


Fig. 50. Temperature (panel (a)) and entropy (panel (b)) versus energy ε for the softening parameter value $\delta = 10^{-5}$. Four values of the energy, indicated by the short-dashed vertical lines, can be identified from this picture: $\varepsilon_1 \simeq -93$ and $\varepsilon_2 \simeq 6$ bound from below and above the region of inequivalence of ensembles. $\varepsilon_t \simeq 0$ is the first order phase transition energy in the microcanonical ensemble. $\varepsilon_a \simeq -66$ limits from below the negative specific heat region. $\beta_t^{-1} \simeq 15$, represented with a dashed line in panel (a), is the canonical transition temperature. The solid lines represent the analytical solutions of temperature and entropy in the high energy phase. They are extended slightly below ε_t, in the metastable phase. The insets in panels (a) and (b) show a zoom of temperature and entropy around ε_t, revealing a temperature jump.

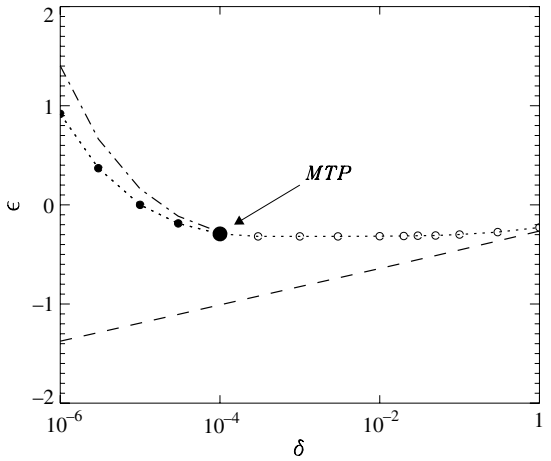


Fig. 51. Phase diagram for the Self-Gravitating Ring model, see Eqs. (436) and (437). The dash–dotted line represents the upper energy of the metastable low energy phase, the dashed line is instead the lower energy of the metastable high energy phase. They bound the region of phase coexistence. The dotted line joining the points is the phase transition line, which changes order at the microcanonical tricritical point (MTP). The filled (open) circles belong to the first (second) order phase transition line.

dipolar interactions offer a relatively simple way to observe a direct manifestation of long-range interactions in samples of suitable shape.

Generally, Heisenberg exchange interaction between electronic spins is much higher (usually a few order of magnitude) than magnetic dipolar energy, and therefore the only role played by the dipolar interaction is the introduction of anisotropies of the spontaneous magnetization. On the other hand, magnetic ordering in nuclear spins, governed by dipolar interaction alone, would require experiments performed at nanokelvin temperatures [286]. Here, we describe an example in which the dipolar interaction between electronic spins acquires, through the careful preparation of the sample, an effective strength in a suitable range of temperatures, such that its long-range character gives rise to an interesting dynamical effect related to a phase transition [287].

Compounds of the type $(C_\nu H_{2\nu+1}NH_3)_2CuCl_4$ have been considered in Refs. [288–290], where the interest was the observation of intrinsic localized modes (discrete breathers). These compounds, organized in a face centered orthorombic crystal, are layered spin structures, in which the weak magnetic interlayer interaction is antiferromagnetic for $\nu > 1$ and ferromagnetic for $\nu = 1$. The relevant variables are the $s = 1/2$ spins of the Cu^{2+} ions, that are placed in two-dimensional layers. The hard axis of the magnetic interaction is orthogonal to the layers, that therefore are the easy planes. However, also the

in-plane interaction is anisotropic: an easy and a "second easy" axis can be determined. The cited works focused on the antiferromagnetic case $\nu = 2$. Here, we are interested in the ferromagnetic case $\nu = 1$, i.e. to the compound $(C_1H_3NH_3)_2CuCl_4$, called bis(Methylammonium) tetrachloro-copper [291].

When writing the Hamiltonian of the system, we adopt a coordinate frame in which the hard axis is the x-axis, while the z-axis and the y-axis are the easy and "second easy" axis belonging to the two-dimensional layers, respectively. The Hamiltonian reads therefore

$$H = -W \sum_{I,\langle i,j\rangle} \left(s_{Ii}^z s_{Ij}^z + \eta s_{Ii}^y s_{Ij}^y + \xi s_{Ii}^x s_{Ij}^x\right) - w \sum_{I,\langle i,j\rangle} \vec{s}_{Ii} \cdot \vec{s}_{I+1,j} + \sum_{Ii\neq Jj} \frac{2\mu_B^2}{r^3} \left(\vec{s}_{Ii} \cdot \vec{s}_{Jj} - 3\frac{(\vec{s}_{Ii} \cdot \vec{r})(\vec{s}_{Jj} \cdot \vec{r})}{r^2}\right), \quad (439)$$

where the first, second and third sums represent the intralayer exchange interaction, the interlayer exchange interaction and the dipolar interaction, respectively. The capital indices number the layers, while the lowercase indices denote the spins. In the first sum $\langle i, j\rangle$ refers to nearest-neighbors within the same layer and in the second sum to nearest-neighbors in adjacent layers; the last sum is extended over all pairs of spins of the system. The intralayer exchange constant W is much larger than the interlayer one w, for example $W \simeq 10^4 w$ [292]. The parameters ξ and η, both smaller than 1 are related to the out-of-plane and in-plane anisotropies. In the dipolar interaction, μ_B is Bohr's magneton and r is the modulus of $\vec{r}$, the vector between the sites of the spins $\vec{s}_{Ii}$ and $\vec{s}_{Jj}$.

The large W/w ratio determines the existence, for a given shape of the sample, of a temperature range in which all the spins of a single layer are ferromagnetically ordered, such that the spin vector $\vec{s}_{Ii}$ can be considered as independent of the index i, $\vec{s}_I \equiv \vec{s}_{Ii}$, while full three-dimensional ordering has not yet been reached. This situation arises if the temperature is below the single layer ordering temperature, of the order of W, and if in addition the number n of spins in each single layer is such that $nw < W$. Under such conditions the thermodynamic and dynamical properties of the system will be determined by the interlayer exchange constant w and by the long-range dipolar interaction. Exploiting the fact that all spins in a layer are ordered, the Hamiltonian can be cast in a one-dimensional form. To this purpose, one has to use a procedure employed in the treatment of dipolar forces, in which the short-range contribution (here including the interaction between nearest-neighbors) and the long-range contribution are treated separately [77]. The latter contribution gives rise to shape-dependent terms. Here, we consider a rod-shaped sample, with short sides along the out-of-plane x-axis and the in-plane y-axis, and long sides along the in-plane z-axis. With this shape, the long-range dipolar interaction produces an interaction term in the Hamiltonian which is typical of an ellipsoidal shape [77,293]. The magnetization in the sample varies only along the x-axis. If we denote by N the number of layers of the sample, the Hamiltonian is then transformed to [293,294]

$$H_N = n\left[B_x \sum_{J=1}^{N} \left(s_J^x\right)^2 + B_y \sum_{J=1}^{N} \left(s_J^y\right)^2 - 2\omega_{ex} \sum_{J=1}^{N-1} \left(s_J^y s_{J+1}^y + s_J^z s_{J+1}^z\right) - \frac{\omega_M}{N}\left(\sum_{J=1}^{N} s_J^z\right)^2 + \frac{\omega_M}{2N}\left(\sum_{J=1}^{N} s_J^y\right)^2\right]. \quad (440)$$

The first two sums come from the first sum in Eq. (439), and describe the intralayer exchange interaction. Here $B_x = 4W(1-\xi)$ and $B_y = 4W(1-\eta)$. Since for this compound the out-of-plane anisotropy is much larger than the in-plane one, we have that $B_x \gg B_y$; a constant additive term involving W has been neglected. The third sum in Eq. (440) is a combination of the interlayer exchange interaction (second sum in (439)) and the nearest-neighbor interlayer dipolar interaction, while the intralayer nearest-neighbor dipolar interaction produces a constant term that has been neglected. In this sum, ω_{ex} is given by

$$\omega_{ex} = 2w - \frac{2\mu_B^2}{r_b^3}\left(2 - \frac{3r_a^2}{2r_b^2}\right), \quad (441)$$

where r_a and r_b are the distances between nearest-neighbor spins within the same layers and adjacent layers, respectively. The last two mean-field terms in Eq. (440) come from the long-range part of the dipolar interaction. This contribution is proportional to a combination of the square of the components of the magnetization density vector, with the coefficients of the combination depending on sample shape. In this case, where the magnetization varies along the x-axis, one has to consider the average magnetization density vector [293]. In these mean-field terms, ω_M is given by $\omega_M = (4\pi/3)(2\mu_B^2/v_0)$, where v_0 is the volume of the unit cell of the lattice. In the effective Hamiltonian (440), we have neglected all terms including the x component of the spins which emerge from dipolar or exchange interlayer forces, since B_x is much larger than B_y, ω_M and ω_{ex}.

Hamiltonian (440) governs the dynamics of the vector spin $\vec{s}_J$ through the torque equation

$$\hbar \frac{d\vec{s}_J}{dt} = \vec{s}_J \times \vec{\mathcal{H}}_J, \quad (442)$$

where $\vec{\mathcal{H}}_J = -\partial H_N/\partial \vec{s}_J$ is an effective magnetic field acting on the spin $\vec{s}_J$. The numerical results that we are going to present shortly are based on the integration of the last equation. To make a connection with the analytical results, we introduce an approximation in the Hamiltonian. We define $\vec{S}_J \equiv 2\vec{s}_J$, $S_J^y = \sqrt{1-(S_J^x)^2}\sin\theta_J$ and $S_J^z = \sqrt{1-(S_J^x)^2}\cos\theta_J$,

which yields the equations of motion in terms of the angular variables θ_j and S_j^x. Noting that B_x is much larger than all the other parameters, which implies that $\sqrt{1-(S_j^x)^2} \approx 1$, the equations of motion are simplified as follows

$$\hbar \frac{\mathrm{d}\theta_j}{\mathrm{d}t} = -nB_x S_j^x. \tag{443}$$

This dynamics is consistent with the following effective Hamiltonian, which includes only the angles θ_j's, and their time derivatives

$$\begin{aligned} H_N' &= \frac{2H_N}{n\omega_M} \\ &= \frac{1}{2}\sum_{j=1}^{N}\left[\frac{\mathrm{d}\theta_j}{\mathrm{d}t}\right]^2 - \frac{\omega_{ex}}{\omega_M}\sum_{j=1}^{N-1}\cos(\theta_{j+1}-\theta_j) + \frac{B_y}{2\omega_M}\sum_{j=1}^{N}\sin^2\theta_j - \frac{1}{2N}\left(\sum_{j=1}^{N}\cos\theta_j\right)^2 + \frac{1}{4N}\left(\sum_{j=1}^{N}\sin\theta_j\right)^2, \end{aligned} \tag{444}$$

where we have introduced a dimensionless time via the transformation $t \to tn\sqrt{B_x\omega_M}/\hbar$. We note the similarity of this Hamiltonian with the Hamiltonian (384) of the chain of rotators coupled by a nearest-neighbor interaction and a mean-field one studied in Section 6.1.2. With respect to Hamiltonian (384), the one given in Eq. (444) is not invariant under global rotations, because of the presence of the term proportional to B_y and of the difference between the coefficients of the two mean-field terms. However, if $B_y > 0$, which is the case experimentally, these differences do not change qualitatively [271] the phase diagram obtained in Section 6.1.2.

Let us now discuss the dynamical effect that could in principle be observed in experiments with this type of sample. We have shown that the XY model with nearest-neighbor and mean-field interactions has, in particular, a second order ferromagnetic phase transition for all positive values of the nearest-neighbor coupling. We have also shown that, below the critical energy, there is another energy threshold, below which ergodicity is broken, with an inaccessible range of values of the magnetization around 0 (contrary to the case $K > 0$ of Section 6.1.2, the inaccessibility of a magnetization range around 0 has to be considered a breaking of ergodicity, since now the lack of invariance under global rotation can forbid the passage from $m_x > 0$ to $m_x < 0$). For the particular values of the parameters in Hamiltonian (444), one finds [287] that the critical energy is $\varepsilon_c \simeq 0.376$, and the threshold below which there is breaking of ergodicity is $\varepsilon = -0.3$. Besides that, because of the lack of global rotational invariance, the spontaneous magnetization has only two possible directions, i.e., $\theta_j = 0$ and $\theta_j = \pi$. If we consider the dynamics of such a system, at an energy between the threshold for ergodicity breaking and the critical energy, and with a finite number of degrees of freedom, we should therefore observe that the modulus of the magnetization will fluctuate around that of the spontaneous magnetization, and from time to time will flip between the two directions of the magnetization. We expect that the flips will be more and more rare as the system size increases and approaching the ergodicity breaking threshold [160]. The sudden flips are due to the fact that, contrary to short-range systems, the formation of domains with different directions of the magnetization is not possible.

Numerical experiments have been performed using Hamiltonian (440) and the torque equation (442) [287]. We here present results concerning the two energies $\varepsilon = 0.2$ and $\varepsilon = 0.4$ (in the adimensional values computed after the transformation to Hamiltonian (444)). The first value is below the critical energy but still far enough from the ergodicity breaking threshold to observe magnetization flips on a reasonable time scale. The second value is above the critical energy, and the system should fluctuate around zero magnetization. In Fig. 52, we show the results of the numerical simulations that confirm the expectations.

From the numerical data, it is also possible to obtain the probability distribution function for the magnetization, shown in the insets of Fig. 52 by the open circles. The logarithm of the probability distributions obtained numerically is compared with that computed analytically at the corresponding energy values on the basis of Hamiltonian (444). A comparison is then made between the numerical distribution functions obtained by direct simulations of Hamiltonian (440) and the expression $\tilde{s}(\varepsilon, x)$ given by Eq. (391), obtained for Hamiltonian (444). The agreement is good although not perfect. The effect which leads to disagreement is twofold. First, one should have obtained the expression of $\tilde{s}(\varepsilon, x)$ for Hamiltonian (440), which is in principle possible. Second, we have taken x as playing the role of the spontaneous magnetization m. However, we have already emphasized several times that the parameter x coming from the Hubbard–Stratonovich transformation coincides with the spontaneous magnetization only at equilibrium. The reasonably good agreement that we obtain justifies a posteriori this identification, which is valid only close to equilibrium.

In summary, our main purpose here was to see that a system which is realizable in the laboratory can be used to observe the peculiarities predicted for long-range interacting systems. We have emphasized the importance of sample shape. For instance: for spherical or approximately spherical, samples, spin flips will not be observed, since the mean-field terms in the Hamiltonian will not introduce any anisotropy.

7. Conclusions

7.1. Summary

It should be now clear that the dynamics and thermodynamics of systems with long-range interactions is a rich and fascinating field. It has been shown how, changing the range of the interactions, deeply renew several aspects of statistical

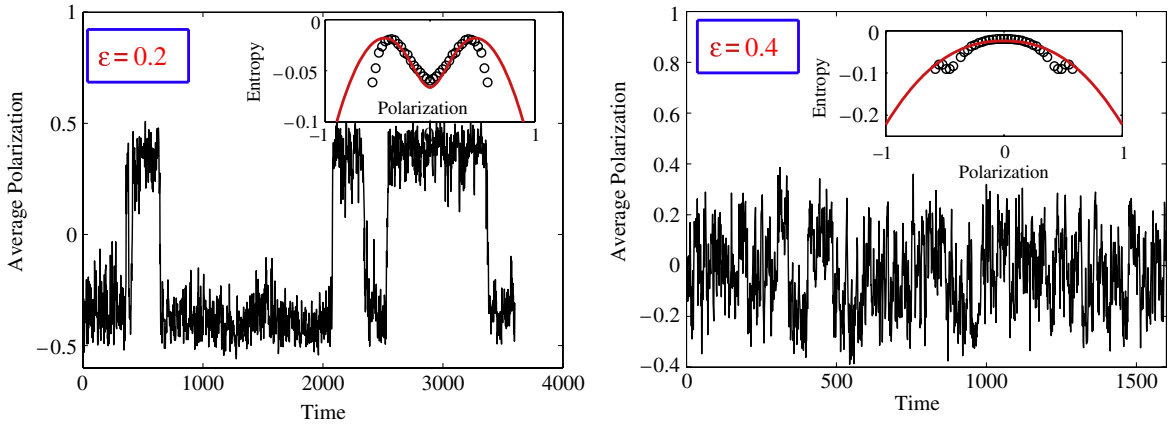


Fig. 52. Numerical simulations of the one dimensional spin chain model, Eqs. (440) and (442), with two different energies: $\varepsilon = 0.2$ (left panel) and $\varepsilon = 0.4$ (right panel), i.e. below and above the second order phase transition energy $\varepsilon_c \simeq 0.376$ predicted for Hamiltonian (444). The number of layers in the simulations is $N = 100$. The insets show the corresponding entropy curves (solid lines) obtained for Hamiltonian (444) and the data obtained from the probability distribution function of the magnetization (open circles) for Hamiltonian (440).

Table 1
Summary of the important properties of the different models studied in this review. C for continuous and D for discrete. Y for Yes, N for No, while the slot is left empty when the answer is not known. L + S means long-range plus short-range interactions.

Model	Variable	Ensemble inequivalence	Negative specific heat	Negative susceptibility	Ergodicity breaking	Computable entropy	Section
BEG	D	Y	Y		Y	Y	4.2
3 states Potts	D	Y	Y	N	N	Y	4.5.2
Ising L + S	D	Y	Y		Y	Y	6.1.1
α-Ising	D	Y	N			Y	6.2.1
HMF	C	N	N			Y	4.4
XY L + S	C	Y	Y		Y	Y	6.1.2
α-HMF	C	N	N				6.2.2
Generalized HMF	C	Y	Y		Y	Y	4.5.4
Mean-field ϕ^4	C	Y	N	Y	Y	Y	4.5.5
Colson–Bonifacio	C	N	N	N	N	Y	4.5.6
SGR	C	Y	Y	N	Y	Y	6.2.3

mechanics. We have encountered many unusual properties: ensemble inequivalence, negative specific heat and negative susceptibility, ergodicity breaking, quasi-stationary states, algebraic relaxation and anomalous diffusion. Rather than being contradictory with standard statistical mechanics, they extend its domain of validity. It should be remarked that the analysis of statistical ensembles for long-range systems privileges the use of microcanonical ensemble, since all canonical macrostates can be realized in the microcanonical ensemble, but the converse is not true: this is the essence of ensemble inequivalence. These properties are of course not found in all models and for any value of the thermodynamic parameters. Table 1 summarizes the main features of the models studied in this review.

We have mostly restricted our analysis to mean-field models. However, we have shown that the properties that we have discussed extend also to non-mean-field interactions, as presented in Section 6. Recently, new analytical methods have been developed to solve models with long-range interactions at equilibrium: they all aim at the direct derivation of the entropy. We have in particular presented in full detail the large deviation method in Section 4.5. Relaxation to equilibrium is a key topic in systems with long-range interactions, which historically began with the realization that self-gravitating systems cannot converge to a statistical equilibrium. By considering a much simpler model, that nevertheless is representative of long-range interacting systems, we have shown that relaxation to equilibrium proceeds in two steps. On a fast time scale, one observes relaxation to quasi-stationary states that are successfully described using a theoretical approach proposed by Lynden-Bell. On a longer time scale, one indeed observes relaxation to Boltzmann–Gibbs equilibrium. In connection with this complex relaxation to equilibrium, we have extensively discussed kinetic equations for systems with long-range interactions (Klimontovich, Vlasov, Lenard–Balescu, Fokker–Planck). All this is presented in Section 5.

7.2. Open problems

Before closing this review, we would like to briefly comment on a few topics that we have discussed but which would need a wider treatment since it is clear that the present review opens a "Pandora's box".

Another source of negative specific heat which has not discussed at all in this review is thermodynamic instability. Even if a system is short-range, attractive and the interaction is regularized at short distances, energy density can be unbounded

from below and both potential energy and kinetic energy will increase as a consequence of the formation of a cluster. This phenomenon has been discussed in a series of papers beginning with Refs. [295,296].

For two-dimensional and geophysical flows, Chavanis and Sommeria [297], as well as Bouchet and Venaille [68], obtain different equilibrium states depending on the shape of the domain. Following a classification of phase transitions previously derived by Barré and Bouchet [42], they exhibit the first example of bicritical points and second order azeotropy for long-range interacting systems. There is a close connection between two-dimensional fluid dynamics and two-dimensional electron plasmas. In this latter context, experiments have shown the relaxation towards different quasi stationary states [74,75,298], and recently theoretical frameworks have been proposed [255,299]. Electron plasmas could also constitute experimental setups to observe in the future the new type of phase transitions predicted in Ref. [42].

While the initial violent relaxation has been successfully characterized, the slow collisional process that leads to the final equilibrium is still an open field of research. Chavanis has carefully investigated this problem and has proposed a rich kinetic theory approach [205]. As an application, collisional relaxation has been numerically investigated for the HMF model in Ref. [222].

Concerning self-gravitating systems, one should quote the remarkable observation of a sequence of transient states in the long-term stellar dynamical evolution. Density distributions of such states are well fitted by stellar polytropes [300,301]. Similar observations have been reported for the HMF model in Ref. [222]. Recently, the problem of the robustness of quasi-stationary states with respect to noise has been investigated [302]. Gravitational clustering in finite and infinite systems has been numerically investigated by Joyce, Sylos Labini and collaborators [303]. These authors have tried to assess if simulations performed with a finite number of particles reproduce the mean-field Vlasov limit and to understand the universal features of halo structures.

An interesting direction of investigation is the one followed by Barré and Gonçalves, that considers models on random graphs. Using the large deviation cavity method, analytical solutions of the Potts spins systems on a random k-regular graph have been recently derived in both the canonical and microcanonical ensembles. The analytical solution, confirmed by numerical Metropolis and Creutz simulations, clearly demonstrate the presence of a region with negative specific heat and, consequently, ensemble inequivalence between canonical and microcanonical ensembles [304].

An alternative method has been used to describe quasi-stationary states. These states, according to the method of non-extensive statistics proposed by Tsallis [224], should be maximum entropy states of a generalized Boltzmann entropy defined as $S_q = k_B\,(1-\sum_i p_i^q)/(q-1)$, where p_i is the normalized probability of state i. As correctly pointed out by Tsallis and coworkers [305,306], the presence of quasi-stationary states is tightly linked to the non-commutability of the infinite time limit with the thermodynamic limit for systems with long-range interactions. Velocity distributions that maximize Tsallis entropy S_q are q-Gaussian defined by $G_q(x) = A(1-(1-q)\eta x^2)^{1/1-q}$, where A is a normalization constant and η controls the width of the distribution. These distributions do not convincingly fit numerical data of the HMF model [305,306]. Moreover, no theoretical approach has been proposed which leads to the derivation of such distributions for a specific case, like the HMF model. For this reason, this theory has not shown any predictive power in this field.

Although unsuccessful to explain quasi-stationary states, Tsallis approach has led to the introduction of interesting models and concepts. A relevant model, that we also discuss in Section 6.2.2, is the α-HMF model, which has been introduced in Ref. [276] to discuss the behavior of the maximal Lyapunov exponent in systems with long-range interactions: this is still an open and interesting question. Algebraic decay of time correlations and anomalous diffusion within the HMF model were discovered by Rapisarda and coworkers [233,307–309]. However, the approach to this problem developed in Sections 5.2.2 and 5.2.3 is able to explain all numerical observations without resorting to non-extensive statistics. Moreover, although motivated by non-extensive statistics, the work in Ref. [310] concludes that quasi-stationary states, originally looked at by Rapisarda and coworkers, do not show anomalous diffusion, as also previously reported in Ref. [247].

7.3. Perspectives

The study of long-range interactions has certainly overcome the pioneering era where it has been important to become acquainted with basic facts and concepts that looked odd initially, like negative specific heat. Having passed to a more mature period, this field, besides strengthening its foundations, should incorporate methods and tools developed in other disciplines and look towards the realizations of well devised experiments that could lead to the observations of the effects predicted by the theory.

From the mathematical point of view, one is faced to interesting and difficult problems for what concerns both equilibrium statistical mechanics and out-of-equilibrium dynamics. It would be important to develop large deviation methods in order to allow the treatment of systems with slowly decaying interactions as was done for instance for the α-Ising model [127] and of systems with both short and long-range interactions. For what concerns the dynamical properties, several significant advances have been recently obtained in kinetic theory concerning the study of the solutions of Vlasov–Poisson and Vlasov–Maxwell systems. Both the difficulties of incorporating in the theory the singularity of the Coulomb or gravitational potential and the enlargement of the treatable initial conditions have been addressed (see Ref. [311]).

When going to systems where the interaction is long-range, but not mean-field, the difficulty of performing numerical simulations is related to the fact that algorithms are intrinsically of order N^2 for each time step, where N is the number of particles. Because of this, for gravitational systems, dedicated computers have been built (e.g. GRAPE in Japan [280])

and sophisticated numerical codes have been developed [312]. The issue of devising better algorithms is central for the improvements of the simulations and has not yet been seriously faced.

We have contributed with this review to convey the idea that many of the effects that were previously attributed to gravitational systems appear also in other systems, as was early stressed in two-dimensional hydrodynamics [7,24]. We have already mentioned in Section 3 two-dimensional hydrodynamics, two-dimensional elasticity, charged and dipolar systems, small systems. We would like to concentrate here on three physical examples that are more likely to lead in the near future to experimental realizations: the free electron laser, cold atoms and lattice dipolar systems.

The cooperative effect leading to coherent emission of radiation in the Free Electron Laser (FEL) has been described in Section 4.5.6, using an analogy with phase transitions. In a realistic setting, convergence to Boltzmann–Gibbs equilibrium will never occur because of the extremely elevated number of electrons, compared to the length of the wiggler. That is why it is highly probable that FELs will always operate in regions of parameters dominated by quasi-stationary states. As we know from the analysis of such states in the HMF model, they show phase transitions from homogeneous to inhomogeneous macroscopic distributions (see Section 5.2.4). Such transitions, together with the characterization of quasi-stationary states, could be the object of an experimental study with linear FELs [126].

Cold atomic systems offer a promising laboratory for testing the predictions of the statistical mechanics of systems with long-range interactions. These systems range from ionized atoms confined by an external potential which interact via unscreened Coulomb repulsion, to neutral atoms that interact by means of Van der Waals type forces. Additionally, dipolar gases have been recently the subject of extensive experimental studies [313]. Initially, some attention had been paid to the possible realization of true gravitational-like classical interaction induced by external laser fields [314]. However, this hope has failed until now due to technical difficulties. When the wave-like behavior of matter can be neglected, leading to a semi-classical approximation, and laser light is far from resonance, a fully Hamiltonian (microcanonical) treatment of Bose–Einstein condensates is viable. Moreover, ensemble inequivalence for bosonic gases is reminiscent of the ensemble inequivalence encountered for long-range interacting systems [315–317]. However, it has been later realized that this type of ensemble inequivalence is of a different nature and should be analyzed in itself. A potentially fruitful seed has recently emerged in the study of cold atoms in optical high-Q cavities. For these systems, the interaction between two atoms is independent of the distance, because they are coupled via the back-scattering of photons within the cavity. The interaction can lead to collective instabilities and self-organization phenomena, known as CARL [8,311,318]. Models of the CARL phenomenon have strong similarities with those of FEL. Therefore, we expect that the study of phase transitions and of quasi-stationary states could be realized also for these laboratory systems.

We finally argue that dipolar condensed matter systems should be seriously considered to check some features of long-range interactions. Indeed, it has been remarked since long time that composites like $HoRhB_4$, where dipolar interactions dominate over Heisenberg exchange interactions, have a mean-field critical behavior [319]. More recently, such results have been extended to diluted models which simulate the effect of disorder, making the Sherrington–Kirpatrick spin-glass behavior experimentally accessible [320]. The importance of long-range dipolar interactions has been also recognized for pyrochlore lattice structures, that show the strongly degenerate and frustrated "spin ice" phase [81].

In the final section of this review, we have briefly mentioned experimental applications, because we think that theoretical results obtained in the study of systems with long-range interactions have nowadays reached a level of understanding that might allow, in the near future, the experimental verifications of curious effects like negative specific heat, ergodicity breaking, quasi-stationary states.

Acknowledgements

We would like to warmly thank our collaborators Angel Alastuey, Mickael Antoni, Andrea Antoniazzi, Romain Bachelard, Julien Barré, Fausto Borgonovi, Freddy Bouchet, Francesco Califano, Gian-Luca Celardo, Cristel Chandre, Pierre-Henri Chavanis, Giovanni De Ninno, Yves Elskens, Duccio Fanelli, Marie-Christine Firpo, Andrea Giansanti, Dieter Gross, Alessio Guarino, Haye Hinrichsen, Peter Holdsworth, Ramaz Khomeriki, Tetsuro Konishi, Hiroko Koyama, Raman Johal, Vito Latora, Xavier Leoncini, Stefano Lepri, François Leyvraz, Leonardo Lori, Gianluca Morelli, Daniele Moroni, David Mukamel, Andrea Rapisarda, Nir Schreiber, Luca Sguanci, Takayuki Tatekawa, Alessandro Torcini, Hugo Touchette, Romain Vasseur, Yoshi Yamaguchi. We thank for financial support the Galileo program of the French–Italian University "Study and control of models with a large number of interacting particles" and the COFIN07-PRIN program "Statistical physics of strongly correlated systems at and out of equilibrium" of the Italian MIUR. Finally, SR and AC thank the ENS Lyon for hospitality and financial support, and SR thanks Université de Provence for hospitality and financial support.

Appendix A. Proof of min–max inequality

Let us consider the function of two variables $f(x, y)$. Under quite general conditions on f, the following inequality holds

$$\sup_x \inf_y f(x, y) \leq \inf_y \sup_x f(x, y). \tag{445}$$

Let us give a sketchy proof of this property, which is used in the text to prove that microcanical entropy (160) is always smaller or equal than canonical entropy (158). Let us denote by (x_1, y_1) the extremum that satisfies the l.h.s. of (445), then

$$f(x_1, y_1) \leq f(x_1, y), \quad \forall y. \tag{446}$$

Similarly, the extremum that satisfies the r.h.s., denoted (x_2, y_2) satisfies

$$f(x_2, y_2) \geq f(x, y_2), \quad \forall x. \tag{447}$$

Then,

$$f(x_1, y_1) \leq f(x_1, y_2) \leq f(x_2, y_2), \tag{448}$$

which proves the inequality (445).

Appendix B. Evaluation of the Laplace integral outside the analyticity strip

The microcanonical partition function in formula (162) can be expressed, using the Laplace representation of the Dirac δ function as follows

$$\Omega(\varepsilon, N) = \frac{1}{2\pi \mathrm{i}} \int_{\beta - \mathrm{i}\infty}^{\beta + \mathrm{i}\infty} \mathrm{d}\lambda \mathrm{e}^{N\lambda\varepsilon} Z(\lambda, N), \tag{449}$$

with $\beta > 0$; this is the expression given in (164). As explained in Section 4.4.4, we divide the integral in (449) in three intervals, defined by $\lambda_I < -\delta$, $-\delta < \lambda_I < \delta$ and $\lambda_I > \delta$, respectively, with $0 < \delta < \Delta$. Here we show that the contribution to the integral in λ coming from values of λ_I outside the strip, i.e. for values of λ_I with $|\lambda_I| > \Delta$, is exponentially small in N.

Let us then consider first the value of $Z(\lambda, N)$ in the two external intervals, i.e., for $|\lambda_I| > \Delta$. We have

$$Z(\beta + \mathrm{i}\lambda_I, N) = \sum_{\{S_1, \ldots, S_N\}} \exp\left\{-\beta H(\{S_i\})\right\} \exp\left\{-\mathrm{i}\lambda_I H(\{S_i\})\right\}. \tag{450}$$

We see that this expression is proportional to the canonical expectation value $\langle \exp(-\mathrm{i}\lambda_I H)\rangle$, that we expect to be exponentially small for large N. We confirm this expectation rewriting the last expression as

$$Z(\beta + \mathrm{i}\lambda_I, N) = N \int \mathrm{d}\varepsilon \exp\left(-N\left[\beta\varepsilon + \mathrm{i}\lambda_I\varepsilon - s(\varepsilon)\right]\right), \tag{451}$$

where we have used the expression of the canonical partition in terms of the microcanonical entropy as in (61). This is an integral with a large phase. For N going to infinity, its value will be determined by the value of the integrand for ε equal to the integration extremes and to the values of the possible nonanalyticities of $s(\varepsilon)$, all denoted by ε_k (see, e.g., Ref. [321]). We then have

$$Z(\beta + \mathrm{i}\lambda_I, N) \sim N \sum_k c_k \exp\left(-N\left[\beta\varepsilon_k + \mathrm{i}\lambda_I\varepsilon_k - s(\varepsilon_k)\right]\right), \tag{452}$$

where c_k are coefficients that could in principle be determined. It is clear that the successive integration over λ_I in any one of the two external intervals of integration will then give a vanishing contribution, due to the very large oscillations. We are then left with

$$\Omega(\varepsilon, N) \overset{N\to+\infty}{\sim} \frac{1}{2\pi \mathrm{i}} \int_{\beta - \mathrm{i}\delta}^{\beta + \mathrm{i}\delta} \mathrm{d}\lambda\, \mathrm{e}^{N\lambda\varepsilon} Z(\lambda, N), \tag{453}$$

which is the first equality given in Eq. (165).

Appendix C. Differentiability of the function $\bar{\phi}(\lambda_1, \ldots, \lambda_n)$

We give a proof that the function $\bar{\phi}(\lambda_1, \ldots, \lambda_n)$, defined in Eq. (193), is differentiable, for real values of the λ_i, when the functions $\mu_k(x)$ are given by sums of one-particle functions

$$\mu_k(x) = \frac{1}{N} \sum_{i=1}^{N} g_k(q_i, p_i) \quad k = 1, \ldots, n. \tag{454}$$

In this case, the canonical partition function is given by

$$\bar{Z}(\lambda_1, \ldots, \lambda_n) = \left(\int \mathrm{d}p\mathrm{d}q \exp\left[-\sum_{k=1}^{n} \lambda_k g_k(q, p) \right] \right)^N. \tag{455}$$

Therefore we have

$$\bar{\phi}(\lambda_1, \ldots, \lambda_n) = -\ln\left(\int \mathrm{d}p\mathrm{d}q \exp\left[-\sum_{k=1}^{n} \lambda_k g_k(q, p) \right] \right). \tag{456}$$

In the range of λ_i where the integral in Eq. (456) is defined, $\tilde{\phi}$ is continuous. Besides, differentiating with respect to λ_i we have:

$$\frac{\partial\tilde{\phi}}{\partial\lambda_i}(\lambda_1,\ldots,\lambda_n)=\frac{\int \mathrm{d}p\mathrm{d}q\, g_i(q,p)\exp\left[-\sum_{k=1}^{n}\lambda_k g_k(q,p)\right]}{\int \mathrm{d}p\mathrm{d}q\exp\left[-\sum_{k=1}^{n}\lambda_k g_k(q,p)\right]}. \tag{457}$$

Also this function is continuous, under the only hypothesis that the expectation value of the observable $g_i(q,p)$ is finite for all the allowed values of the λ_i in the canonical ensemble defined by $\bar{Z}(\lambda_1,\ldots,\lambda_n)$.

Appendix D. Autocorrelation of the fluctuations of the one-particle density

Using the definition of the Fourier transform and formula (271), one gets

$$\left\langle \delta f\,(k,p,0)\,\delta f\left(k',p',0\right)\right\rangle = \int_0^{2\pi}\frac{\mathrm{d}\theta}{2\pi}\int_0^{2\pi}\frac{\mathrm{d}\theta'}{2\pi}\,\mathrm{e}^{-\mathrm{i}(k\theta+k'\theta')}\left\langle \delta f\,(\theta,p,0)\,\delta f\left(\theta',p',0\right)\right\rangle \tag{458}$$

$$= \int_0^{2\pi}\frac{\mathrm{d}\theta}{2\pi}\int_0^{2\pi}\frac{\mathrm{d}\theta'}{2\pi}\mathrm{e}^{-\mathrm{i}(k\theta+k'\theta')}N\left[\left\langle f_d\,(\theta,p,0)\,f_d\left(\theta',p',0\right)\right\rangle - f_0\,(p)\,f_0\left(p'\right)\right]. \tag{459}$$

The expression of the discrete density function (261) leads then to

$$\left\langle f_d\,(\theta,p,0)\,f_d\left(\theta',p',0\right)\right\rangle = \frac{1}{N^2}\Bigg\langle\sum_{j=1}^{N}\delta\left(\theta-\theta_j\right)\delta\left(p-p_j\right)\delta\left(\theta-\theta'\right)\delta\left(p-p'\right)$$

$$+\sum_{i\neq j}\delta\left(\theta-\theta_j\right)\delta\left(p-p_j\right)\delta\left(\theta'-\theta_i\right)\delta\left(p'-p_i\right)\Bigg\rangle \tag{460}$$

$$= \frac{1}{N^2}\left[N\left\langle f_d\,(\theta,p,0)\right\rangle\delta\left(\theta-\theta'\right)\delta\left(p-p'\right)+N(N-1)f_2(0,\theta,p,\theta',p')\right] \tag{461}$$

$$= \frac{1}{N}f_0\,(p)\,\delta\left(\theta-\theta'\right)\delta\left(p-p'\right)+f_0\,(p)\,f_0\left(p'\right)+h_2(\theta,p,\theta',p',0), \tag{462}$$

where we used the following definition of the correlation function h_2

$$f_2(\theta,p,\theta',p',0) = \left\langle\delta\left(\theta-\theta_j\right)\delta\left(p-p_j\right)\delta\left(\theta'-\theta_i\right)\delta\left(p'-p_i\right)\right\rangle \tag{463}$$

$$= \frac{N}{N-1}\left[f_0\,(p)\,f_0\left(p'\right)+h_2(\theta,p,\theta',p')\right]. \tag{464}$$

Substituting expression (462) in Eq. (459), we find

$$\left\langle \delta f\,(k,p,0)\,\delta f\left(k',p',0\right)\right\rangle = \int_0^{2\pi}\frac{\mathrm{d}\theta}{2\pi}\frac{f_0\,(p)}{2\pi}\mathrm{e}^{-\mathrm{i}(k+k')\theta}\delta\left(p-p'\right)+\int_0^{2\pi}\frac{\mathrm{d}\theta}{2\pi}\int_0^{2\pi}\frac{\mathrm{d}\theta'}{2\pi}\,N\mathrm{e}^{-\mathrm{i}(k\theta+k'\theta')}h_2(\theta,p,\theta',p') \tag{465}$$

$$= \frac{f_0\,(p)}{2\pi}\delta_{k,-k'}\delta\left(p-p'\right)+\frac{1}{2\pi}\delta_{k,-k'}\,\mu(k,p,p') \tag{466}$$

$$= \frac{\delta_{k,-k'}}{2\pi}\left[f_0(p)\delta(p-p')+\mu(k,p,p')\right]. \tag{467}$$

In the passage from (465) to (466), we have used the fact that h_2 depends only on the difference $\theta-\theta'$. Besides, it decays rapidly to zero in a range $(\theta-\theta')\sim 1/N$, so that $\mu(k,p,p')$ is of order 1.

Appendix E. Derivation of the Fokker–Planck coefficients

As mentioned in Section 5.2.2, the derivation of the Fokker–Planck equation (328) is performed in the time range $1\ll t\ll N$. Within this approximation, expressions (329) and (330) are thus replaced by

$$A(p,t)=\frac{1}{t}\langle(p(t)-p(0))\rangle_{p(0)=p} \tag{468}$$

$$B(p,t)=\frac{1}{t}\langle(p(t)-p(0))^2\rangle_{p(0)=p}. \tag{469}$$

Looking at Eq. (327), we see that, at the order $1/N$, we need to compute the following averages

$$\frac{1}{\sqrt{N}}\left\langle \frac{\partial \delta v}{\partial \theta}(\theta(0)+p(0)t,t)\right\rangle \tag{470}$$

and

$$\frac{1}{N}\left\langle \frac{\partial^2 \delta v}{\partial \theta^2}(\theta(0)+p(0)t_1,t_1)\frac{\partial \delta v}{\partial \theta}(\theta(0)+p(0)t_2,t_2)\right\rangle \tag{471}$$

for the first moment, and

$$\frac{1}{N}\left\langle \frac{\partial \delta v}{\partial \theta}(\theta(0)+p(0)t_1,t_1)\frac{\partial \delta v}{\partial \theta}(\theta(0)+p(0)t_2,t_2)\right\rangle \tag{472}$$

for the second one. The fact that the position and the momentum of the test-particle at time 0 are given is specified inside the dependence on θ of the derivatives of δv. The condition of given initial angle and momentum of the test-particle produces corrections of the averages in Eqs. (470)–(472), with respect to their unconditioned values. These corrections being of order $1/\sqrt{N}$, one gets that at order $1/N$ they have to be taken into account only for Eq. (470). Note that the unconditioned value of that equation vanishes.

Let us start with Eq. (472). To compute it, it is necessary to rewrite Eq. (313) as

$$\left\langle \delta v(k,\omega)\delta v(k',\omega')\right\rangle = 2\pi^3\delta_{k,-k'}\left(\delta_{k,1}+\delta_{k,-1}\right)\frac{\delta(\omega+\omega')}{\left|\tilde{D}(\omega,k)\right|^2}\int \mathrm{d}p' f_0(p')\delta(\omega-p'k). \tag{473}$$

By the inverse Fourier–Laplace transform, we then have

$$\left\langle \frac{\partial \delta v}{\partial \theta}(\theta(0)+pt_1,t_1)\frac{\partial \delta v}{\partial \theta}(\theta(0)+pt_2,t_2)\right\rangle$$

$$= \frac{1}{(2\pi)^2}\sum_{k=-\infty}^{+\infty}\sum_{k'=-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega\int_{-\infty}^{+\infty}\mathrm{d}\omega'\,(-kk')\,\mathrm{e}^{\mathrm{i}k(\theta(0)+pt_1)-\mathrm{i}\omega t_1}\mathrm{e}^{\mathrm{i}k'(\theta(0)+pt_2)-\mathrm{i}\omega' t_2}\left\langle \delta v(k,\omega)\delta v(k',\omega')\right\rangle \tag{474}$$

$$= \frac{\pi}{2}\sum_{k=-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega\, k^2\mathrm{e}^{\mathrm{i}(kp-\omega)(t_1-t_2)}\frac{\delta_{k,1}+\delta_{k,-1}}{\left|\tilde{D}(\omega,k)\right|^2}\int \mathrm{d}p' f_0(p')\delta(\omega-p'k) \tag{475}$$

$$= \frac{\pi}{2}\int_{-\infty}^{+\infty}\mathrm{d}\omega\,\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\left[\mathrm{e}^{\mathrm{i}(p-\omega)(t_1-t_2)}+\mathrm{e}^{-\mathrm{i}(p-\omega)(t_1-t_2)}\right]. \tag{476}$$

Eq. (327) shows that the second moment appearing in the right-hand side of Eq. (469) is determined by the integral of the last expression in t_1 and t_2 from 0 to t. More precisely

$$B(p,t) = \frac{1}{t}\frac{1}{N}\frac{\pi}{2}\int_0^t \mathrm{d}t_1\int_0^t \mathrm{d}t_2\int_{-\infty}^{+\infty}\mathrm{d}\omega\,\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\left[\mathrm{e}^{\mathrm{i}(p-\omega)(t_1-t_2)}+\mathrm{e}^{-\mathrm{i}(p-\omega)(t_1-t_2)}\right] \tag{477}$$

$$= \frac{1}{t}\frac{1}{N}\pi\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\int_0^t \mathrm{d}s\int_0^{t-s}\mathrm{d}t_2\left[\mathrm{e}^{\mathrm{i}(p-\omega)s}+\mathrm{e}^{-\mathrm{i}(p-\omega)s}\right] \tag{478}$$

$$= \frac{1}{t}\frac{1}{N}\pi \int_{-\infty}^{+\infty} \mathrm{d}\omega \frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2} \int_0^t \mathrm{d}s\,(t-s)\left[\mathrm{e}^{\mathrm{i}(p-\omega)s} + \mathrm{e}^{-\mathrm{i}(p-\omega)s}\right] \tag{479}$$

$$\overset{t\to+\infty}{\sim} \frac{1}{N} 2\pi \int_{-\infty}^{+\infty} \mathrm{d}\omega \frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2} \pi\delta(p-\omega) \tag{480}$$

$$= \frac{1}{N} 2\pi^2 \frac{f_0(p)}{\left|\tilde{D}(p,1)\right|^2} = \frac{2}{N} D(p), \tag{481}$$

which is the expression (332).

For what concerns Eq. (471), starting again from Eq. (473), we obtain

$$\left\langle \frac{\partial^2 \delta v}{\partial\theta^2}(\theta(0)+pt_1,t_1)\frac{\partial \delta v}{\partial\theta}(\theta(0)+pt_2,t_2)\right\rangle$$

$$= \frac{1}{(2\pi)^2}\sum_{k=-\infty}^{+\infty}\sum_{k'=-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega\int_{-\infty}^{+\infty}\mathrm{d}\omega'\,(-\mathrm{i}k^2k')\,\mathrm{e}^{\mathrm{i}k(\theta(0)+pt_1)-\mathrm{i}\omega t_1}\mathrm{e}^{\mathrm{i}k'(\theta(0)+pt_2)-\mathrm{i}\omega' t_2}\left\langle \delta v(k,\omega)\delta v(k',\omega')\right\rangle \tag{482}$$

$$= \mathrm{i}\frac{\pi}{2}\sum_{k=-\infty}^{+\infty}\int_{-\infty}^{+\infty}\mathrm{d}\omega\,k^3\,\mathrm{e}^{\mathrm{i}(kp-\omega)(t_1-t_2)}\frac{\delta_{k,1}+\delta_{k,-1}}{\left|\tilde{D}(\omega,k)\right|^2}\int \mathrm{d}p' f_0(p')\delta(\omega-p'k) \tag{483}$$

$$= \mathrm{i}\frac{\pi}{2}\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\left[\mathrm{e}^{\mathrm{i}(p-\omega)(t_1-t_2)} - \mathrm{e}^{-\mathrm{i}(p-\omega)(t_1-t_2)}\right]. \tag{484}$$

Eq. (327) shows that the contribution $A_1(p,t)$ to the first moment appearing in the right-hand side of Eq. (468) is given by

$$A_1(p,t) = \frac{1}{t}\int_0^t \mathrm{d}u\int_0^u \mathrm{d}u_1\int_0^{u_1}\mathrm{d}u_2\left\langle \frac{\partial^2\delta v}{\partial\theta^2}(\theta(0)+pu,u)\frac{\partial\delta v}{\partial\theta}(\theta(0)+pu_2,u_2)\right\rangle \tag{485}$$

$$= \frac{1}{t}\frac{1}{N}\frac{\mathrm{i}\pi}{2}\int_0^t \mathrm{d}u\int_0^u \mathrm{d}u_1\int_0^{u_1}\mathrm{d}u_2\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\left[\mathrm{e}^{\mathrm{i}(p-\omega)(u-u_2)} - \mathrm{e}^{-\mathrm{i}(p-\omega)(u-u_2)}\right] \tag{486}$$

$$= -\frac{1}{t}\frac{1}{N}\pi\,\mathrm{Im}\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\int_0^t\mathrm{d}u\int_0^u\mathrm{d}u_2\,(u-u_2)\mathrm{e}^{\mathrm{i}(p-\omega)(u-u_2)} \tag{487}$$

$$= -\frac{1}{t}\frac{1}{N}\pi\,\mathrm{Re}\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\frac{\partial}{\partial\omega}\int_0^t\mathrm{d}u\int_0^u\mathrm{d}u_2\mathrm{e}^{\mathrm{i}(p-\omega)(u-u_2)} \tag{488}$$

$$= \frac{1}{t}\frac{1}{N}\pi\,\mathrm{Re}\int_{-\infty}^{+\infty}\mathrm{d}\omega\left(\frac{\partial}{\partial\omega}\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\right)\int_0^t\mathrm{d}u\frac{\mathrm{i}}{p-\omega}\left[1-\mathrm{e}^{\mathrm{i}(p-\omega)u-u_2}\right] \tag{489}$$

$$= \frac{1}{t}\frac{1}{N}\pi\int_{-\infty}^{+\infty}\mathrm{d}\omega\left(\frac{\partial}{\partial\omega}\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\right)\int_0^t\mathrm{d}u\frac{\sin(p-\omega)u}{p-\omega} \tag{490}$$

$$= \frac{1}{t}\frac{1}{N}2\pi\int_{-\infty}^{+\infty}\mathrm{d}\omega\left(\frac{\partial}{\partial\omega}\frac{f_0(\omega)}{\left|\tilde{D}(\omega,1)\right|^2}\right)\frac{1}{(p-\omega)^2}\sin^2\frac{(p-\omega)}{2} \tag{491}$$

$$\underset{\sim}{t\to+\infty}\ \frac{1}{t}\frac{1}{N}2\pi\left(\frac{\partial}{\partial p}\frac{f_0(p)}{\left|\tilde{D}(p,1)\right|^2}\right)\int_{-\infty}^{+\infty}\mathrm{d}\omega\,\frac{1}{(p-\omega)^2}\sin^2\frac{(p-\omega)}{2} \tag{492}$$

$$=\quad \frac{1}{N}\pi^2\left(\frac{\partial}{\partial p}\frac{f_0(p)}{\left|\tilde{D}(p,1)\right|^2}\right)=\frac{1}{N}\frac{\mathrm{d}}{\mathrm{d}p}D(p). \tag{493}$$

For what concerns Eq. (470), let us rewrite Eq. (305) as

$$\widetilde{\delta v}(k,\omega)=-\frac{\pi\left(\delta_{k,1}+\delta_{k,-1}\right)}{\tilde{D}(\omega,k)}\int_{-\infty}^{+\infty}\mathrm{d}p'\,\frac{\widehat{\delta f}(k,p',0)}{\mathrm{i}(p'k-\omega)} \tag{494}$$

where we have neglected the $k=0$, that will not contribute. The condition on given initial angle and momentum of the test particle implies that

$$\left\langle\widehat{\delta f}(k,p',0)\right\rangle=\frac{1}{\sqrt{N}}\frac{1}{2\pi}\mathrm{e}^{-\mathrm{i}k\theta(0)}\delta(p-p'). \tag{495}$$

We therefore have

$$\left\langle\frac{\partial\delta v}{\partial\theta}(\theta(0)+pu,u)\right\rangle=-\frac{1}{\sqrt{N}}\frac{1}{(2\pi)^2}\sum_{k=-\infty}^{+\infty}\int_{\mathcal{C}}\mathrm{d}\omega\,(\mathrm{i}k)\mathrm{e}^{\mathrm{i}k(\theta(0)+pu)-\mathrm{i}\omega u}\frac{\pi\left(\delta_{k,1}+\delta_{k,-1}\right)}{\tilde{D}(\omega,k)}\frac{\mathrm{e}^{-\mathrm{i}k\theta(0)}}{\mathrm{i}(pk-\omega)} \tag{496}$$

$$=-\frac{1}{\sqrt{N}}\frac{1}{4\pi}\int_{\mathcal{C}}\mathrm{d}\omega\left[\frac{\mathrm{e}^{\mathrm{i}(p-\omega)u}}{(p-\omega)\tilde{D}(\omega,1)}+\frac{\mathrm{e}^{-\mathrm{i}(p+\omega)u}}{(p+\omega)\tilde{D}(\omega,-1)}\right] \tag{497}$$

$$=-\frac{1}{\sqrt{N}}\frac{1}{4\pi}\left(\mathcal{P}\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{1}{p-\omega}\left[\frac{\mathrm{e}^{\mathrm{i}(p-\omega)u}}{\tilde{D}(\omega,1)}+\frac{\mathrm{e}^{-\mathrm{i}(p-\omega)u}}{\tilde{D}(-\omega,-1)}\right]+\mathrm{i}\pi\left[\frac{1}{\tilde{D}(p,1)}-\frac{1}{\tilde{D}(-p,-1)}\right]\right) \tag{498}$$

$$=-\frac{1}{\sqrt{N}}\frac{1}{4\pi}\left(2\,\mathrm{Re}\,\mathcal{P}\int_{-\infty}^{+\infty}\mathrm{d}\omega\,\frac{1}{p-\omega}\frac{1}{\left|\tilde{D}(\omega,1)\right|^2}\left[\tilde{D}^*(\omega,1)\mathrm{e}^{\mathrm{i}(p-\omega)u}\right]+2\pi^3\frac{f_0'(p)}{\left|\tilde{D}(p,1)\right|^2}\right). \tag{499}$$

From Eq. (327), one sees that the contribution $A_2(p,t)$ of this term to the first moment appearing in the right hand side of Eq. (468) is given by

$$A_2(p,t)=\quad\frac{1}{t}\frac{1}{N}\frac{1}{4\pi}\int_0^t\mathrm{d}u\left(2\,\mathrm{Re}\,\mathcal{P}\int_{-\infty}^{+\infty}\mathrm{d}\omega\frac{1}{p-\omega}\frac{1}{\left|\tilde{D}(\omega,1)\right|^2}\left[\tilde{D}^*(\omega,1)\mathrm{e}^{\mathrm{i}(p-\omega)u}\right]+2\pi^3\frac{f_0'(p)}{\left|\tilde{D}(p,1)\right|^2}\right) \tag{500}$$

$$=\quad\frac{1}{N}\frac{1}{2}\pi^2\frac{f_0'(p)}{\left|\tilde{D}(p,1)\right|^2}+\frac{1}{t}\frac{1}{N}\frac{1}{4\pi}2\,\mathrm{Re}\,\mathcal{P}\int_{-\infty}^{+\infty}\mathrm{d}\omega\,\frac{1}{(p-\omega)^2}\frac{\tilde{D}^*(\omega,1)}{\left|\tilde{D}(\omega,1)\right|^2}\left[2\mathrm{i}\sin^2\frac{p-\omega}{2}+\sin(p-\omega)t\right] \tag{501}$$

$$\underset{\sim}{t\to+\infty}\ \frac{1}{N}\frac{1}{2}\pi^2\frac{f_0'(p)}{\left|\tilde{D}(p,1)\right|^2}+\frac{1}{N}\frac{1}{2}\,\mathrm{Re}\,\frac{\mathrm{i}\tilde{D}^*(\omega,1)}{\left|\tilde{D}(\omega,1)\right|^2} \tag{502}$$

$$=\quad\frac{1}{N}\frac{1}{2}\pi^2\frac{f_0'(p)}{\left|\tilde{D}(p,1)\right|^2}+\frac{1}{N}\frac{1}{2}\pi^2\frac{f_0'(p)}{\left|\tilde{D}(p,1)\right|^2}=\frac{1}{N}\pi^2\frac{f_0'(p)}{\left|\tilde{D}(p,1)\right|^2}=\frac{1}{N}\frac{1}{f_0}\frac{\partial f_0}{\partial p}D(p). \tag{503}$$

Adding Eq. (493) for $A_1(p,t)$ and (503) for $A_2(p,t)$ we obtain Eq. (331) for $A(p,t)$.

References

[1] T. Dauxois, S. Ruffo, E. Arimondo, M. Wilkens (Eds.), Dynamics and Thermodynamics of Systems with Long-Range Interactions, in: Lecture Notes in Physics, vol. 602, Springer, 2002.
[2] A. Campa, A. Giansanti, G. Morigi, F. Sylos Labini, Dynamics and Thermodynamics of systems with long-range interactions: Theory and Experiment, AIP Conference Proceedings 970 (2008).

[3] T. Padmanabhan, Physics Reports 188 (1990) 285.
[4] P.H. Chavanis, International Journal of Modern Physics B 20 (2006) 3113.
[5] Y. Elskens, D. Escande, Microscopic Dynamics of Plasmas and Chaos, IOP Publishing, Bristol, 2002.
[6] R. Robert, Comptes Rendus de l'Académie des Sciences: Série I Maths 311 (1990) 575.
[7] J. Miller, Physical Review Letters 65 (1990) 2137.
[8] S. Slama, G. Krenz, S. Bux, C. Zimmermann, P.W. Courteille, Scattering in a High-Q Ring Cavity, in [2].
[9] Ph. Chomaz, F. Gulminelli, Phase transitions in finite systems, in [1].
[10] K. Huang, Statistical Mechanics, John Wiley and Sons Ed., 1987.
[11] W. Thirring, Zeitschrift für Physik 235 (1970) 339.
[12] P. Hertel, W. Thirring, Annals of Physics 63 (1971) 520.
[13] M. Kiessling, J.L. Lebowitz, Letters in Mathematical Physics 42 (1997) 43.
[14] J. Barré, D. Mukamel, S. Ruffo, Physical Review Letters 87 (2001) 030601.
[15] R. Emden, Gaskugeln, Teubner, Leipzig, 1907.
[16] A.S. Eddington, The Internal Constitution of Stars, Cambridge University Press, 1926.
[17] V.A. Antonov, Vest. Leningrad Gros. Univ. 7 (1962) 135.
[18] D. Lynden-Bell, R. Wood, Monthly Notices of the Royal Astronomical Society 138 (1968) 495.
[19] D.H.E. Gross, Microcanonical Thermodynamics, World Scientific, Singapore, 2001.
[20] R.S. Ellis, H. Touchette, B. Turkington, Physica A 335 (2004) 518.
[21] R. Balian, From Microphysics to Macrophysics: Methods and Applications of Statistical Physics, Springer-Verlag, Berlin, 1992.
[22] J. Binney, S. Tremaine, Galactic Dynamics, in: Princeton Series in Astrophysics, 1987.
[23] D. Lynden-Bell, Monthly Notices of the Royal Astronomical Society 136 (1967) 101.
[24] P.H. Chavanis, J. Sommeria, R. Robert, The Astrophysical Journal 471 (1996) 385.
[25] V. Latora, A. Rapisarda, S. Ruffo, Physical Review Letters 83 (1999) 2104.
[26] Y.Y. Yamaguchi, J Barré, F. Bouchet, T. Dauxois, S. Ruffo, Physica A 337 (2004) 36.
[27] S. Fauve, Pattern forming instabilities, in: C. Godrèche, P. Manneville (Eds.), Hydrodynamics and Nonlinear Instabilities, Cambridge University Press, 1995.
[28] T. Dauxois, M. Peyrard, Physics of Solitons, Cambridge University Press, 2006.
[29] D. Ruelle, Statistical Mechanics: Rigorous Results, Benjamin, New York, 1969.
[30] M. Kac, G.E. Uhlenbeck, P.C. Hemmer, Journal of Mathematical Physics 4 (1963) 216.
[31] D.H.E. Gross, Thermo-Statistics or Topology of the Microcanonical Entropy Surface, in [1].
[32] M.K.H. Kiessling, Journal of Statistical Physics 55 (1989) 203.
[33] T. Padmanabhan, Statistical mechanics of gravitating systems in static and expanding backgrounds, in [1].
[34] P.H. Chavanis, Statistical mechanics of two-dimensional vortices and stellar systems, in [1].
[35] F. Baldovin, E. Orlandini, Physical Review Letters 96 (2006) 240602.
[36] F. Baldovin, E. Orlandini, Physical Review Letters 97 (2006) 100601.
[37] H.A. Posch, W. Thirring, Physical Review E 74 (2006) 051103.
[38] D. Lynden-Bell, R.M. Lynden-Bell, Europhysics Letters 82 (2008) 43001.
[39] A. Ramirez-Hernandez, H. Larralde, F. Leyvraz, Physical Review Letters 100 (2008) 120601.
[40] A. Ramirez-Hernandez, H. Larralde, F. Leyvraz, Physical Review E 78 (2008) 061133.
[41] L. Velazquez, S. Curilef, Journal of Physics A 42 (2009) 095006.
[42] F. Bouchet, J. Barré, Journal of Statistical Physics 118 (2005) 1073.
[43] J. Villain, Reflets de la Physique 7 (2008) 10.
[44] P. Hertel, W. Thirring, Communications in Mathematical Physics 24 (1971) 22.
[45] L.-L. Chabanol, F. Corson, Y. Pomeau, Europhysics Letters 50 (2000) 148.
[46] J.C. Maxwell, Cambridge Philosophical Society's Transactions, vol. XII, 1876, p. 90.
[47] D. Lynden-Bell, Physica A 263 (1999) 29.
[48] B. Castaing, An Introduction to hydrodynamics, in: C. Godrèche, P. Manneville (Eds.), Hydrodynamics and Nonlinear Instabilities, Cambridge University Press, 1995.
[49] G.L. Eyink, K.R. Sreenivasan, Review of Modern Physics 78 (2006) 87.
[50] L. Onsager, Nuovo Cimento Supplement 6 (1949) 279.
[51] A. Alastuey, M. Magro, P. Pujol, Physique et Outils Mathématiques: Méthodes et Exemples, CNRS Editions and EDP Sciences, 2008.
[52] C. Marchioro, M. Pulvirenti, Mathematical Theory of Incompressible Nonviscous Fluids, Springer-Verlag, New York, Heidelberg, 1994.
[53] E.M. Purcell, R.V. Pound, Physical Review 81 (1951) 279.
[54] G. Joyce, D. Montgomery, Journal of Plasma Physics 10 (1973) 107–121.
[55] T.S. Lundgren, Y.B. Pointin, Journal of Statistical Physics 17 (1977) 232.
[56] R. Robert, J. Sommeria, Journal of Fluid Mechanics 229 (1991) 291.
[57] J. Michel, R. Robert, Communications in Mathematical Physics 159 (1994) 195.
[58] B. Turkington, A. Majda, K. Haven, M. DiBattista, The Proceedings of the National Academy of Sciences (USA) 98 (2001) 12346.
[59] F. Bouchet, Mécanique Statistique pour des Écoulements Géophysiques, Ph.D. Thesis, Université Joseph Fourier, Grenoble, 2001.
[60] F. Bouchet, J. Sommeria, Journal of Fluid Mechanics 464 (2002) 165.
[61] P.H. Chavanis, J. Sommeria, Physical Review E 65 (2002) 026302.
[62] E. Caglioti, P.L. Lions, C. Marchioro, M. Pulvirenti, Communications in Mathematical Physics 174 (1995) 229.
[63] R.S. Ellis, K. Haven, B. Turkington, Journal of Statistical Physics 101 (2000) 999.
[64] R.S. Ellis, K. Haven, B. Turkington, Nonlinearity 15 (2002) 239.
[65] J. Sommeria, Two-dimensional Turbulence, in: M. Lesieur, A. Yaglom, F. David (Eds.), New Trends in Turbulence, vol. 74, Les Houches Summer School, 2001, p. 385.
[66] P. Tabeling, Physics Reports 362 (2002) 1.
[67] H.G.H. Clercx, G.J.F. van Heijst, Applied Mechanics Reviews 62 (2009) 020802.
[68] A. Venaille, F. Bouchet, Physical Review Letters 102 (2009) 104501.
[69] N.I. Muskhelishvili, Some Basic Problems of the Mathematical Theory of Elasticity, P. Noordhoff, Groningen, 1953.
[70] J.L. Lebowitz, E.H. Lieb, Physical Review Letters 22 (1969) 613.
[71] E.H. Lieb, J.L. Lebowitz, Advances in Mathematics 9 (1972) 316.
[72] D.C. Brydges, P.A. Martin, Journal of Statistical Physics 96 (1999) 1163.
[73] D.H. Dubin, T.M. O'Neil, Review of Modern Physics 71 (1999) 87.
[74] X.-P. Huang, C.F. Driscoll, Physical Review Letters 72 (1994) 2187.
[75] K.S. Fine, A.C. Cass, W.G. Flynn, C.F. Driscoll, Physical Review Letters 75 (1995) 3277.
[76] M.K.H. Kiessling, T. Neukirch, The Proceedings of the National Academy of Sciences (USA) 100 (2003) 1510.
[77] L.D. Landau, E.M. Lifshitz, Course of Theoretical Physics. T. 8: Electrodynamics of Continuous Media, 1984.
[78] R.B. Griffiths, Physical Review 176 (1968) 655.
[79] S. Banerjee, R.B. Griffiths, M. Widom, Journal of Statistical Physics 93 (2004) 109.
[80] C. Kittel, Physical Review 82 (1951) 965.

[81] S.T. Bramwell, M.J.P. Gingras, Science 294 (2001) 1495.
[82] B. Barbara, Private communication, 2008.
[83] D.H.E. Gross, Physics Reports 279 (1997) 119.
[84] Ph. Chomaz, F. Gulminelli, Nuclear Physics A 153 (1999) 647.
[85] Ph. Chomaz, F. Gulminelli, V. Duflot, Physical Review Letters 64 (2001) 046114.
[86] Ph. Chomaz, F. Gulminelli, European Physical Journal A 30 (2006) 317.
[87] Ph. Chomaz, Phase transitions in finite systems using information theory, in [2].
[88] F. Gulminelli, Ph. Chomaz, Physical Review E 66 (2002) 046108.
[89] P. Labastie, R.L. Whetten, Physical Review Letters 65 (1990) 1567.
[90] D. Lynden-Bell, R.M. Lynden-Bell, Monthly Notices of the Royal Astronomical Society 181 (1977) 405.
[91] M. Schmidt, R. Kusche, T. Hippler, J. Donges, W. Kronmüller, B. von Issendorff, H. Haberland, Physical Review Letters 86 (2001) 1191.
[92] F. Gobet, B. Farizon, M. Farizon, M.J. Gaillard, J.P. Buchet, M. Carré, T.D. Märk, Physical Review Letters 87 (2001) 203401.
[93] M. D'Agostino, F. Gulminelli, P. Chomaz, M. Bruno, F. Cannata, R. Bougault, F. Gramegna, I. Iori, N. Le Neindre, G.V. Margagliotti, A. Moroni, G. Vannini, Physics Letters B 473 (2000) 219.
[94] M. Belkacem, V. Latora, A. Bonasera, Physical Review C 52 (1995) 271.
[95] M. Farizon, B. Farizon, S. Ouaskit, T.D. Märk, Fragment size distributions and caloric curve in collision induced cluster fragmentation, in [2].
[96] D. Frenkel, B. Smit, Understanding Molecular Simulation: From Algorithm to Applications, Academic Press, 1996.
[97] H.H. Rugh, Physical Review Letters 78 (1997) 772.
[98] L. Van Hove, Physica 15 (1949) 951.
[99] H. Touchette, Physics Reports 478 (2009) 1.
[100] R.M. Lynden-Bell, in: O. Lahav, E. Terlevich, R.J. Terlevich (Eds.), Gravitational Dynamics, Cambridge University Presss, Cambridge, 1996.
[101] R.M. Lynden-Bell, Molecular Physics 86 (1995) 1353.
[102] K. Binder, Physica A 319 (2003) 99.
[103] L.G. MacDowell, P. Virnau, M. Müller, K. Binder, Journal of Chemical Physics 120 (2004) 5293.
[104] M. Antoni, A. Torcini, Physical Review E 57 (1998) R6233.
[105] A. Torcini, M. Antoni, Physical Review E 59 (1999) 2746.
[106] M. Antoni, S. Ruffo, A. Torcini, Physical Review E 66 (2002) 025103.
[107] M. Blume, V.J. Emery, R.B. Griffiths, Physical Review A 4 (1971) 1071.
[108] A. Ayuela, N.H. March, Physics Letters A 372 (2008) 5617.
[109] M. Blume, Physical Review 141 (1966) 517.
[110] H.W. Capel, Physica 32 (1966) 966.
[111] J. Katz, Monthly Notices of the Royal Astronomical Society 183 (1978) 765.
[112] E.B. Aronson, C.J. Hansen, Astrophysical Journal 177 (1972) 145.
[113] P.H. Chavanis, Physical Review E 65 (2002) 056123.
[114] J. Barré, D. Mukamel, S. Ruffo, Ensemble inequivalence in mean-field models of magnetisms, in Ref. [1].
[115] B. Stahl, M.K.H. Kiessling, K. Schindler, Planetary and Space Sciences 43 (1995) 271–282.
[116] A. Campa, S. Ruffo, H. Touchette, Physica A 369 (2007) 517.
[117] G. Horwitz, J. Katz, Astrophysical Journal 211 (1977) 226.
[118] G. Horwitz, J. Katz, Astrophysical Journal 222 (1978) 941.
[119] W.C. Saslaw, Gravitational Physics of Stellar and Galactic Systems, Cambridge University Press, 1985.
[120] V.P. Youngkins, B.N. Miller, Physical Review E 62 (2000) 4583.
[121] M. Hénon, Annales d'Astrophysique 27 (1964) 83.
[122] F. Hohl, M.R. Feix, Astrophysical Journal 147 (1967) 1164.
[123] G.B. Rybicki, Astrophysical and Space Sciences 14 (1971) 56.
[124] Lj. Milanovic, H.A. Posch, W. Thirring, Physical Review E 57 (1998) 2763.
[125] P.M. Chaikin, T.C. Lubensky, Principles of Condensed Matter Physics, Cambridge University Press, 1995.
[126] J. Barré, T. Dauxois, G. De Ninno, D. Fanelli, S. Ruffo, Physical Review E, Rapid Communication 69 (2004) 045501 (R).
[127] J. Barré, F. Bouchet, T. Dauxois, S. Ruffo, Journal of Statistical Physics 119 (2005) 677.
[128] M. Antoni, S. Ruffo, Physical Review E 52 (1995) 2361.
[129] T. Dauxois, V. Latora, A. Rapisarda, S. Ruffo, A. Torcini, The Hamiltonian Mean Field model: From dynamics to statistical mechanics and back, in Ref. [1].
[130] P.H. Chavanis, J. Vatteville, F. Bouchet, European Physical Journal B 46 (2005) 61.
[131] R. Toral, Journal of Statistical Physics 114 (2004) 1393.
[132] C. Anteneodo, Physica A 342 (2004) 112.
[133] S. Ruffo, Hamiltonian dynamics and phase transitions, in: S. Benkadda, Y. Elskens, F. Doveil (Eds.), Transport and Plasma Physics, World Scientific Singapore, 1994.
[134] S. Inagaki, Progress in Theoretical Physics 90 (1993) 577.
[135] S. Inagaki, T. Konishi, Publications of the Astronomical Society of Japan 45 (1993) 733.
[136] D. Del Castillo-Negrete, Physics Letters A 241 (1998) 99.
[137] D. Del Castillo-Negrete, Physics of Plasmas 5 (1998) 3886.
[138] T. Konishi, K. Kaneko, Journal of Physics A 25 (1992) 6283.
[139] T. Dauxois, P. Holdsworth, S. Ruffo, European Physical Journal B 16 (2000) 659.
[140] J. Barré, T. Dauxois, S. Ruffo, Physica A 295 (2001) 254.
[141] J. Barré, F. Bouchet, T. Dauxois, S. Ruffo, Physical Review Letters 89 (2002) 110601.
[142] J. Barré, F. Bouchet, T. Dauxois, S. Ruffo, European Physical Journal B 29 (2002) 577.
[143] D. Jeong, J. Choi, M.Y. Choi, Physical Review E 74 (2006) 0556106.
[144] A. Campa, A. Giansanti, D. Moroni, Journal of Physics A 36 (2003) 6897.
[145] M. Antoni, H. Hinrichsen, S. Ruffo, Chaos, Solitons and Fractals 13 (2002) 393.
[146] L. Velazquez, R. Sospedra, J.C. Castro, F. Guzman, On the dynamical anomalies in the Hamiltonian Mean Field model, cond-mat/0302456.
[147] F. Leyvraz, S. Ruffo, Journal of Physics A: Mathematical and General 35 (2002) 285.
[148] A. Campa, A. Giansanti, Physica A 340 (2004) 170.
[149] R.S. Ellis, Entropy, Large Deviations, and Statistical Mechanics, Springer-Verlag, New-York, 1985.
[150] A. Dembo, O. Zeitouni, Large Deviations Techniques and their Applications, Springer-Verlag, New York, 1998.
[151] R. Robert, Journal of Statistical Physics 65 (1991) 531.
[152] R.S. Ellis, Physica D 133 (1999) 106.
[153] I. Ispolatov, E.G.D. Cohen, Physica A 295 (2001) 475.
[154] D.H.E. Gross, E.V. Votyakov, European Physical Journal B 15 (2000) 115.
[155] P. de Buyl, D. Mukamel, S. Ruffo, AIP Conferences Proceedings 800 (2005) 533.
[156] F. Bouchet, T. Dauxois, D. Mukamel, S. Ruffo, Physical Review E 77 (2008) 011125.
[157] P.A. Lebwohl, G. Lasher, Physical Review A 6 (1972) 426.
[158] F. Borgonovi, G.L. Celardo, M. Maianti, E. Pedersoli, Journal of Statistical Physics 116 (2004) 1435.

[159] F. Borgonovi, G.L. Celardo, A. Musesti, R. Trasarti-Battistoni, P. Vachal, Physical Review E 73 (2006) 039903.
[160] G.L. Celardo, J. Barré, F. Borgonovi, S. Ruffo, Physical Review E 73 (2006) 011108.
[161] D. Mukamel, S. Ruffo, N. Schreiber, Physical Review Letters 95 (2005) 240604.
[162] M. Creutz, Physical Review Letters 50 (1983) 1411.
[163] R. Desai, R. Zwanzig, Journal of Statistical Physics 19 (1978) 1.
[164] T. Dauxois, S. Lepri, S. Ruffo, Communications in Nonlinear Science and Numerical Simulations 8 (2003) 375.
[165] A. Campa, S. Ruffo, Physica A 385 (2007) 233.
[166] I. Hahn, M. Kastner, Physical Review E 72 (2005) 056134.
[167] I. Hahn, M. Kastner, European Physical Journal B 50 (2006) 311.
[168] W.B. Colson, Physics Letters A 59 (1976) 187.
[169] R. Bonifacio, C. Pellegrini, L.M. Narducci, Optical Communications 50 (1984) 373.
[170] G.M. Zaslavsky, V.F. Shabanov, K.S. Aleksandrov, I.P. Aleksandrova, Soviet Physics JETP 45 (1977) 315.
[171] R. Bonifacio, F. Casagrande, G. Cerchioni, L. De Salvo Souza, P. Pierini, N. Piovella, Rivista del Nuovo Cimento 13 (1) (1990).
[172] M.C. Firpo, Y. Elskens, Physical Review Letters 84 (2000) 3318.
[173] M.C. Firpo, Etude dynamique et statistique de l'interaction onde-particule, Ph.D. Thesis, Université de Provence, 1999.
[174] J. Barré, Mécanique statistique et dynamique hors équilibre de systèmes avec intéractions à longues portées, Ph.D. Thesis, ENS Lyon, 2003.
[175] D. Farina, F. Casagrande, U. Colombo, R. Pozzoli, Physical Review E 49 (1994) 1603.
[176] L. Casetti, M. Pettini, E.G.D. Cohen, Physics Reports 337 (2000) 237.
[177] M. Kastner, Review of Modern Physics 80 (2008) 167.
[178] M. Pettini, Geometry and Topology in Hamiltonian Dynamics and Statistical Mechanics, Springer, Berlin, 2007.
[179] M. Kastner, O. Schnetz, Journal of Statistical Physics 122 (2006) 1195.
[180] M. Kastner, Physica A 359 (2006) 447.
[181] L. Casetti, M. Kastner, Physical Review Letters 97 (2006) 100602.
[182] M. Kastner, S. Schreiber, O. Schnetz, Physical Review Letters 99 (2007) 050601.
[183] G. Gallavotti, Statistical Mechanics: A Short Treatise, Springer, Berlin, 1999.
[184] R. Franzosi, M. Pettini, Physical Review Letters 92 (2004) 060601.
[185] R. Franzosi, M. Pettini, Nuclear Physics B 782 (2007) 189.
[186] R. Franzosi, M. Pettini, Nuclear Physics B 782 (2007) 219.
[187] L. Casetti, E.G.D. Cohen, M. Pettini, Physical Review Letters 82 (1999) 4160.
[188] L. Casetti, M. Pettini, E.G.D. Cohen, Journal of Statistical Physics 111 (2003) 1091.
[189] L. Angelani, L. Casetti, M. Pettini, G. Ruocco, F. Zamponi, Europhysics Letters 62 (2003) 775.
[190] L. Angelani, L. Casetti, M. Pettini, G. Ruocco, F. Zamponi, Physical Review E 71 (2005) 036152.
[191] L. Angelani, G. Ruocco, Physical Review E 76 (2007) 051119.
[192] A.C. Ribeiro Teixeira, D.A. Stariolo, Physical Review E 70 (2004) 016113.
[193] S. Risau-Gusman, A.C. Ribeiro-Teixeira, D.A. Stariolo, Physical Review Letters 95 (2005) 145702.
[194] M. Kastner, O. Schnetz, Physical Review Letters 100 (2008) 160601.
[195] L. Boltzmann, Wien, Ber. 66 (1872) 275.
[196] L.D. Landau, Phys. Z. Sowjetunion 10 (1936) 154.
[197] S. Chandrasekhar, Principles of Stellar Dynamics, University of Chicago Press, 1942.
[198] A.A. Vlasov, Journal of Physics (USSR) 9 (1945) 25.
[199] L.D. Landau, Journal of Physics (USSR) 10 (1946) 25.
[200] A. Lenard, Annals of Physics 10 (1960) 390.
[201] R. Balescu, Physics of Fluids 3 (1960) 52.
[202] M.R. Feix, P. Bertrand, Transport Theory and Statistical Physics 34 (2005) 7.
[203] D.H. Dubin, D.Z. Jin, Physics Letters A 284 (2001) 112.
[204] D.H. Dubin, Physics of Plasmas 10 (2003) 1338.
[205] P.H. Chavanis, Dynamics and thermodynamics of systems with long-range interactions: Interpretation of the different functionals, in [2].
[206] H. Neunzert, Neuere qualitative und numerische Methoden in der Plasmaphysik, Skriptum einer Gastvorlesung an der Gesamthochschule Paderborn, 1975.
[207] H. Neunzert, Fluid Dynamics Transactions 8 (1978).
[208] W. Braun, K. Hepp, Communications in Mathematical Physics 56 (1977) 101.
[209] H. Spohn, Large Scale Dynamics of Interacting Particles, Springer, 1991.
[210] R. Balescu, Equilibrium and Nonequilibrium Statistical Mechanics, Wiley, New York, 1975.
[211] C.R. Willis, R.H. Picard, Physical Review A 9 (1974) 1343.
[212] H.E. Kandrup, Astrophysical Journal 244 (1981) 316.
[213] Yu.L. Klimontovich, The Statistical Theory of Non-equilibrium Processes in a Plasma, MIT Press, 1967.
[214] E.M. Lifshitz, L.P. Pitaevskij, Physical Kinetics, Pergamon Press, Oxford, 1981.
[215] D.R. Nicholson, Introduction to Plasma Theory, John Wiley, 1983.
[216] J. Messer, H. Spohn, Journal of Statistical Physics 29 (1982) 561.
[217] C. Pichon, Dynamics of Self-gravitating Disks, Ph.D. Thesis, Cambridge, 1994.
[218] M.Y. Choi, J. Choi, Physical Review Letters 91 (2003) 124101.
[219] J. Choi, M.Y. Choi, Journal of Physics A 38 (2005) 5659.
[220] F. Bouchet, T. Dauxois, Physical Review E 72 (2005) 045103(R).
[221] C. Anteneodo, R. Vallejos, Physica A 344 (2004) 383.
[222] A. Campa, P.H. Chavanis, A. Giansanti, G. Morelli, Physical Review E 78 (2008) 040102(R).
[223] Y.Y. Yamaguchi, F. Bouchet, T. Dauxois, Journal of Statistical Mechanics: Theory and Experiment P01020 (2007).
[224] C. Tsallis, Journal of Statistical Physics 52 (1988) 479.
[225] B.B. Kadomtsev, O.P. Pogutse, Physical Review Letters 25 (1970) 1155.
[226] R. Balescu, Statistical Mechanics of Charged Particles, Interscience, New York, 1963.
[227] P.H. Chavanis, European Physical Journal B 52 (2006) 61.
[228] P.H. Chavanis, Physica A 361 (2006) 55.
[229] P.H. Chavanis, Physica A 361 (2006) 81.
[230] P.H. Chavanis, Physica A 387 (2008) 787.
[231] P.H. Chavanis, Physica A 387 (2008) 1504.
[232] P.H. Chavanis, Physical Review E 64 (2001) 026309.
[233] V. Latora, A. Rapisarda, C. Tsallis, Physical Review E 64 (2001) 056134.
[234] A. Campa, A. Giansanti, G. Morelli, Physical Review E 76 (2007) 041117.
[235] J. Barré, F. Bouchet, T. Dauxois, S. Ruffo, Y.Y. Yamaguchi, Physica A 365 (2006) 177.
[236] E. Caglioti, F. Rousset, Archive for Rational Mechanics and Analysis 190 (2008) 517.
[237] M. Hauray, P.E. Jabin, Archive for Rational Mechanics and Analysis 183 (2007) 489.
[238] M. Antoni, S. Ruffo, A. Torcini, Europhysics Letters 66 (2004) 645.

[239] P.H. Chavanis, Astronomy and Astrophysics 432 (2005) 117.
[240] N.G. Van Kampen, Stochastic Processes in Physics and Chemistry, North-Holland, Elsevier, 1992.
[241] S. Ichimaru, Basic Principles of Plasma Physics, W. A. Benjamin, Inc., Reading, Mass, 1973.
[242] F. Bouchet, Physical Review E 70 (2004) 036113.
[243] S. Marksteiner, K. Ellinger, P. Zoller, Physical Review A 53 (1996) 3409.
[244] J. Farago, Europhysics Letters 52 (2000) 379.
[245] S. Miccichè, Modeling long-range memory with stationary markovian processes, cond-mat/08060722, 2008.
[246] P.H. Chavanis, M. Lemou, European Physical Journal B 59 (2007) 217.
[247] Y.Y. Yamaguchi, Physical Review E 68 (2003) 066210.
[248] I.F. Potapenko, A.V. Bobylev, C.A. de Azevedo, A.S. de Assis, Physical Review E 56 (1997) 7159.
[249] P.H. Chavanis, M. Lemou, Physical Review E 72 (2005) 061106.
[250] A. Antoniazzi, F. Califano, D. Fanelli, S. Ruffo, Physical Review Letters 98 (2007) 150602.
[251] A. Antoniazzi, D. Fanelli, S. Ruffo, Y.Y. Yamaguchi, Physical Review Letters 99 (2007) 040601.
[252] R. Bachelard, C. Chandre, D. Fanelli, X. Leoncini, S. Ruffo, Physical Review Letters (2008).
[253] Y. Levin, R. Pakter, T. Teles, Physical Review Letters 100 (2008) 040604.
[254] Y. Levin, R. Pakter, F.B. Rizzato, Physical Review E 78 (2008) 021130.
[255] H. Brands, P.H. Chavanis, R. Pasmanter, J. Sommeria, Physics of Fluids 11 (1999) 365.
[256] F. Hohl, J.W. Campbell, Astronomical Journal 73 (1968) 611.
[257] L. Cohen, M. Lecar, Bulletin d'Astronomie, 3ème série, Tome III. fasc 2 213 (1968).
[258] M. Hénon, Bulletin Astronomie, 3ème série, Tome III. fasc. 2 213 (1968).
[259] T. Yamashiro, N. Gouda, M. Sakagami, Progress in Theoretical Physics 88 (1992) 269.
[260] Y.Y. Yamaguchi, Physical Review E 78 (2008) 041114.
[261] R. Robert, J. Sommeria, Physical Review Letters 69 (1992) 2776.
[262] A. Antoniazzi, Y. Elskens, D. Fanelli, S. Ruffo, European Physical Journal B 50 (2006) 603.
[263] A. Antoniazzi, D. Fanelli, J. Barré, P.H. Chavanis, T. Dauxois, S. Ruffo, Physical Review E 75 (2007) 011112.
[264] P.H. Chavanis, European Physical Journal B 53 (2006) 487.
[265] P.H. Chavanis, G. De Ninno, D. Fanelli, S. Ruffo, Out of equilibrium phase transitions in mean-field Hamiltonian dynamics, in: C. Chandre, X. Leoncini, G. Zaslavsky (Eds.), Chaos, Complexity and Transport, World Scientific, 2008, p. 3.
[266] J.F. Nagle, Physical Review A 2 (1970) 2124.
[267] M. Kardar, Physical Review B 28 (1983) 244.
[268] M. Kardar, Physical Review Letters 51 (1983) 523.
[269] H.A. Kramers, G.H. Wannier, Physical Review 60 (1941) 252.
[270] A. Campa, A. Giansanti, D. Mukamel, S. Ruffo, Physica A 365 (2006) 120.
[271] L. Lori, Modelli con interazione a lungo e corto raggio con variabili microscopiche continue, Tesi di Laurea, Università di Firenze, 2008.
[272] F.J. Dyson, Communications in Mathematical Physics 12 (1969) 91.
[273] J. Barré, Physica A 305 (2002) 172.
[274] C. Boucher, R.S. Ellis, B. Turkington, Annals of Probability 27 (1999) 297; Annals of Probability 30 (2002) 2113 (erratum).
[275] L. Casetti, M. Kastner, Physica A 384 (2007) 318.
[276] C. Anteneodo, C. Tsallis, Physical Review Letters 80 (1998) 5313.
[277] F. Tamarit, C. Anteneodo, Physical Review Letters 84 (2000) 208.
[278] J. Barré, Private communication, 2009.
[279] D. Heggie, P. Hut, The Gravitational Million-Body Problem, Cambridge University Press, 2003.
[280] A. Kawai, T. Fukushige, J. Makino, M. Taiji, Publications of the Astronomical Society of Japan 52 (2000) 659.
[281] T. Tsuchiya, T. Konishi, N. Gouda, Physical Review E 50 (1994) 2607.
[282] H. Koyama, T. Konishi, Physics Letters A 279 (2001) 226.
[283] M. Hénon, Mémoire de la Société Royale des Sciences de Liege 55 (1967) 243.
[284] Y. Sota, O. Iguchi, M. Morikawa, T. Tatekawa, K. Maeda, Physical Review E 64 (2001) 056133.
[285] T. Tatekawa, F. Bouchet, T. Dauxois, S. Ruffo, Physical Review E 71 (2005) 056111.
[286] A.S. Oja, V. Lounasmaa, Review of Modern Physics 69 (1997) 1.
[287] A. Campa, R. Khomeriki, D. Mukamel, S. Ruffo, Physical Review B 76 (2007) 064415.
[288] M. Sato, A.J. Sievers, Nature 432 (2004) 486.
[289] M. Sato, A.J. Sievers, Physical Review B 71 (2005) 214306.
[290] J.P. Wrubel, M. Sato, A.J. Sievers, Physical Review Letters 95 (2005) 264101.
[291] G. Chapuis, G. Brunisholz, C. Javet, R. Roulet, Inorganic Chemistry 22 (1983) 455.
[292] A. Dupas, K. Le Dang, J.-P. Renard, P. Veillet, Journal of Physics C (Solid State Physics) 10 (1977) 3399.
[293] A.I. Akhiezer, V.G. Baryakhtar, S.V. Peletminskij, Spin Waves, North Holland, Amsterdam, 1968.
[294] L.Q. English, M. Sato, A.J. Sievers, Physical Review B 67 (2003) 024403.
[295] A. Compagner, C. Bruin, A. Roelse, Physical Review A 39 (1989) 5989.
[296] H.A. Posch, H. Narnhofer, W. Thirring, Physical Review A 42 (1990) 1880.
[297] P.H. Chavanis, J. Sommeria, Journal of Fluid Mechanics 314 (1996) 267.
[298] Y. Kiwamoto, N. Hashizume, Y. Soga, J. Aoki, Y. Kawai, Physical Review Letters 99 (2007) 115002.
[299] R. Kawahara, H. Nakanishi, Journal of Physical Society of Japan 75 (2006) 054001.
[300] A. Taruya, M. Sakagami, Physical Review Letters 90 (2003) 181101.
[301] A. Taruya, M. Sakagami, Monthly Notices of the Royal Astronomical Society 364 (2005) 990.
[302] F. Baldovin, P.H. Chavanis, E. Orlandini, Physical Review E 79 (2009) 011102.
[303] F. Sylos Labini, Gravitational clustering: An overview, in [2].
[304] J. Barré, B. Gonçalves, Physica A 386 (2007) 212.
[305] M. Gell-Mann, C. Tsallis, Nonextensive Entropy-Interdisciplinary Applications, Oxford University Press, 2004.
[306] V. Latora, A. Rapisarda, C. Tsallis, Physica A 305 (2002) 129.
[307] A. Pluchino, V. Latora, A. Rapisarda, Physica A 69 (2004) 056113.
[308] A. Pluchino, V. Latora, A. Rapisarda, Physica A 340 (2004) 187.
[309] A. Pluchino, V. Latora, A. Rapisarda, Physica D 193 (2004) 315.
[310] L.G. Moyano, C. Anteneodo, Physical Review E 74 (2006) 021118.
[311] T. Dauxois, S. Ruffo, L. Cugliandolo (Eds.), Long-Range Interacting Systems, Oxford University Press, 2009.
[312] S.J. Aarseth, Gravitational N-Body Simulations: Tools and Algorithms, Cambridge University Press, Cambridge, 2003.
[313] G. Morigi, Long-range interactions in cold atomic systems: A foreword, in [2].
[314] D. O'Dell, S. Giovanazzi, G. Kurizki, V.M. Akulin, Physical Review Letters 84 (2000) 5687.
[315] D. Boers, M. Holthaus, Canonical statistics of occupation numbers for ideal and weakly interacting Bose–Einstien condensates, in Ref. [1].
[316] M. Wilkens, F. Illuminati, M. Krämer, Journal of Physics B 33 (2000) L779.
[317] R.M. Ziff, G.E. Uhlenbeck, M. Kac, Physics Reports 32 (1977) 169.
[318] R. Bonifacio, L. De Salvo, Nuclear Instruments and Methods A 341 (1994) 360.
[319] H.R. Ott, G. Keller, W. Odoni, L.D. Woolf, M.B. Maple, D.C. Johnston, H.A. Mook, Physical Review B 25 (1982) 477.
[320] M. Ocio, D. Herisson, Fluctuation dissipation relation in a non stationary system: Experimental investigation in a spin glass, in: J.L. Barrat, M. Feigelman, J. Kurchan, J. Dalibard (Eds.), Les Houches Session LXXVII, 2003, p. 605.
[321] C.M. Bender, S.A. Orszag, Advanced Mathematical Methods for Scientists and Engineers, McGrawHill, New-York, 1978 (Chapter 6).