Eur. Phys. J. A **30**, 317–331 (2006)
DOI 10.1140/epja/i2006-10126-5

THE EUROPEAN PHYSICAL JOURNAL A

The challenges of finite-system statistical mechanics

P. Chomaz[1] and F. Gulminelli[2,a]

[1] GANIL (DSM-CEA/IN2P3-CNRS), Blvd. H. Becquerel, F-14076 Caen cédex, France
[2] LPC (IN2P3-CNRS/Ensicaen et Université), F-14076 Caen cédex, France

Received: 22 March 2006 /

Abstract. In this paper, we review the main challenges associated with the statistical mechanics of finite systems, with a particular emphasis on the present understanding of phase transitions in the framework of information theory. We show that this is a very powerful formalism allowing to treat in a thermodynamically consistent way many difficult problems in the statistical treatment of finite, open, transient and expanding systems. The first point we analyze is the problem of boundary conditions, which in the framework of information theory must also be treated statistically. We recall that the different ensembles do not lead to the same equation of states, in particular in the region of a first-order phase transition, and we stress the fact that different statistical ensembles may be relevant to heavy-ion physics depending upon the actual experimental conditions. Finally, we present a coherent description of first-order phase transitions demonstrating the equivalence between the Yang-Lee theorem, the occurrence of bimodalities in the intensive ensemble and the presence of inverted curvatures of the thermodynamic potential of the extensive ensemble. We stress that this discussion is not restricted to the possible occurrence of negative specific heat, but can also include negative compressibilities and negative susceptibilities, and in fact any curvature anomaly of the thermodynamic potential. Since the relevant entropy surface explored in nuclear multifragmentation is not yet well understood and largely debated in the community, the experimental evidence of new thermodynamic anomalies is one of the important challenges of future heavy-ion experiments.

PACS. 05.20.Gg Classical ensemble theory – 25.70.Pq Multifragment emission and correlations – 64.60.-i General studies of phase transitions – 65.40.Gr Entropy and other thermodynamical quantities

1 Introduction

Finite-systems properties, non-extensive thermodynamics, and phase transitions out of the thermodynamic limit are strongly debated issues in many different fields of physics (see for example [1]). This may be the case of non-saturating forces such as the gravitational [2–5] or the Coulombic forces. The system may be too small, as in the case of clusters and nuclei [6–9]. The physics of finite systems is even more complicated since often they are not only small but also open and transient. This implies that the various concepts of thermodynamics and statistical mechanics [10–13] have to be completed and revisited [1,14–18]. Another contribution to this topical issue deals with some aspects of this question [19] and we address the reader to that paper to have a more complete view of the different formalisms that can be applied. In the present paper, we focus on the information theory approach to statistical mechanics [14,16] and we will show that this is a very powerful formalism that allows to address in a consistent way the statistical mechanics of open systems evolving in time, independent of their interaction range and number of constituents.

After a short summary of the statistical-physics concepts, we will summarize the discussion about ensemble inequivalence. Statistical ensembles are presented in many elementary textbooks as qualitatively equivalent and quantitatively almost identical, because they differ only at the fluctuation level. However, for finite systems it is now well documented in the literature [18,20–23] that two ensembles which put different constraints on the fluctuations of the order parameter lead to qualitatively different equations of state close to a first-order phase transition.

This will lead us to the discussion of phase transitions in finite systems. As an example, when energy is the order parameter, the microcanonical (at fixed energy) heat capacity diverges to become negative while the canonical one (at fixed temperature) remains always positive and finite [24–36]. If the number of particle is the order parameter, it is the chemical susceptibility which is expected to present a negative branch in between two divergences in

---

[a] Member of the Institut Universitaire de France;
e-mail: `gulminelli@lpccaen.in2p3.fr`

the fixed number of particle ensemble (microcanonical or canonical) while in the grand canonical it should remain positive. This difference between ensembles can be of primordial importance for mesoscopic systems undergoing a phase transition. Such systems are now studied in many fields of physics, from Bose condensates [37,38] to the quark-gluon plasma [39,40], from cluster melting [6,41] to nuclear fragmentation [7]. Moreover, such inequivalences may survive at the infinite size limit for systems involving long-range forces such as self-gravitating objects [3–5].

We will then present some basic characteristics of first-order phase transitions in finite systems. In particular, we will summarize the mathematical connections between the Yang-Lee approach [42] through the zeroes of the partition sum, the bimodality of the order parameter distribution in the same ensemble, and the anomalous (inverted) curvature of the thermodynamic potential of the ensemble where the order parameter is fixed [43–45]. The best documented example in the literature is the bimodality of the canonical energy distribution being equivalent to negative microcanonical heat capacity [46–48].

Finally, we will skim over the time evolution problem, stressing the need to take into account time odd constraint in the statistical picture, and then conclude presenting three challenges for the field of the thermodynamic properties of finite systems.

2 Finite systems and statistical mechanics

A largely debated issue in the nuclear-physics community is the applicability of thermodynamical concepts like equilibrium, temperature, pressure etc. to objects as tiny as nuclei. How large must a system be for a temperature to be defined? It is well known that the different statistical ensembles only converge in the thermodynamic limit: out of this limit, what is the physical meaning of thermodynamic quantities —say, temperature— evaluated through different ensembles? is there a “correct” ensemble to be used? We all know finite systems can change state or shape, a typical example being the case of isomerization; how many degrees of freedom do we need in order to call this change of state a phase transition?

Let us consider a system that can exist in two single microstates of different energy (a single spin in a magnetic field, a two-level atom in a bath of radiation…) The system being much smaller than its environment, let us consider the case for which the interaction between system and environment can be neglected and we have no reason to believe that the environment will be in any specific state. Then the distribution of the system microstates is simply given by the number of states of the environment

$$\begin{aligned} p^{(n)} &= W(E_t - e_n)/(W(E_t - e_1) + W(E_t - e_2)) \\ &\propto \exp\left(S(E_t - e_n)\right), \end{aligned} \tag{1}$$

where E_t is the total energy (system + environment) and $S = \log W$ is the (microcanonical) entropy associated with the environment. Since $e_n \ll E_t$, a Taylor expansion of the entropy gives

$$S(E_t - e_n) \approx S(E_t) - e_n \frac{\partial S}{\partial E}(E_t)\,; \quad p^{(n)} \propto \exp\left(-\beta e_n\right), \tag{2}$$

where we have introduced $\beta = \partial S \partial E$, the temperature of the environment.

This very simple textbook exercise gives us a number of interesting informations:

- thermodynamic concepts like temperature can be defined for systems having an arbitrary number of degrees of freedom (the minimum being 2 levels);
- Boltzmann-Gibbs statistics naturally emerges as soon as we observe a limited information constructed from a reduced number of degrees of freedom.

If we now take into account a slightly more complicated system with energy states associated with a degeneracy $w(e)$, the energy distribution will be modified to

$$p(e) = \frac{w(e) W(E_t - e)}{\sum_n w(e_n) W(E_t - e_n)} \approx \frac{w(e) \exp\left(-\beta e\right)}{Z_\beta}, \tag{3}$$

where the canonical approximation is still correct if the system is associated with a much smaller number of degrees of freedom than its environment. Equation (3) gives for instance the energy distribution of a thermometer loosely coupled to an otherwise isolated system. Temperature is defined as the response of the thermometer in the most probable energy state $\overline{e}$; if we maximize the distribution (3) we get, assuming that energy can be treated as a continuous variable,

$$\frac{\partial \log W}{\partial E}|_{E_t - \overline{e}} = \frac{\partial \log w}{\partial E}|_{\overline{e}}\,. \tag{4}$$

We then learn that the quantity shared at the most probable energy partition is the microcanonical temperature. This shows that there is no ambiguity in the definition of temperature (and any other thermodynamic quantity) when dealing with small systems. It is important to note that eq. (3) is not limited to the observation of energy, but can apply to the distribution of any generic observable $A = \langle \hat{A} \rangle$. We can then expect that canonical-like ensembles (*i.e.* ensembles where distributions are given by Boltzmann factors) will arise each time that we are isolating a small number of degrees of freedom from a more complex system.

More generally, we will recall in the next section that a statistical description is in order each time that the system is complex enough to have a large number of microstates associated with a given set of relevant observables. The proper statistical ensemble will then depend on the way the system is prepared. If the relevant observables are recognized, equilibrium is therefore a very generic concept that certainly applies to the output of a heavy-ion collision independent of the reaction time.

3 Statistical physics and information theory

Information-theory–based statistical mechanics provides a very powerful framework to a consistent treatment of the

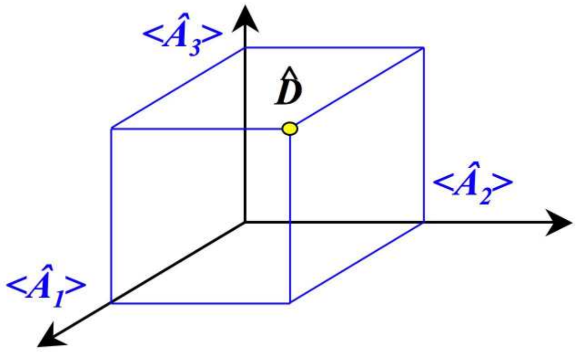

Fig. 1. Illustration of the Liouville space of density matrices: an observation $\langle\hat{A}_i\rangle$ is a projection of $\hat{D}$ on the axis associated with the corresponding observable $\hat{A}_i$.

thermodynamics of finite systems both in the classical and in the quantal world [16,14]. Let us summarize here the essential ingredients. Statistical physics treats statistical ensembles of possible solutions for the considered physical system since the assumption is made that we have only a limited knowledge on it. Such a "macrostate" can be formally be represented by its density matrix

$$\hat{D} = \sum_{(n)} \left|\Psi^{(n)}\right\rangle \, p^{(n)} \, \left\langle\Psi^{(n)}\right|, \tag{5}$$

where the states ("microstates", or "partitions", or "replicas", or simply "events") $|\Psi^{(n)}\rangle$ pertain to the considered Fock or Hilbert space. $p^{(n)}$ is the occurrence probability of the event $|\Psi^{(n)}\rangle$. The result of the measurement of an observable $\hat{A}$ is

$$\langle\hat{A}\rangle_{\hat{D}} = \mathrm{Tr}\,\hat{A}\hat{D}, \tag{6}$$

where Tr means the trace over the quantum Fock or Hilbert space of states $\{|\Psi\rangle\}$.

In the space of Hermitian matrices, the trace provides a scalar product [49,50]

$$\langle\langle\hat{A}||\hat{D}\rangle\rangle = \mathrm{Tr}\,\hat{A}\hat{D}. \tag{7}$$

It is then possible to define an orthonormal basis of Hermitian operators $\{\hat{O}_l\}$ in the observable space, and to interpret the measurement $\langle\hat{O}_l\rangle_{\hat{D}}$ as a coordinate of the density matrix $\hat{D}$ (see fig. 1). The size of the observable space is the square of the dimension of the Hilbert or Fock space, which are in general infinite; therefore in order to describe the system, one is forced to consider a reduced set of (collective) observables $\{\hat{A}_\ell\}$ which are supposed to contain the relevant information. The Gibbs formulation of statistical mechanics can then be derived if the least biased "macrostate" is assumed to be given by the maximization of the entropy[1]

$$S[\hat{D}] = -\,\mathrm{Tr}\,\hat{D}\log\hat{D}, \tag{8}$$

which is nothing but the opposite of the Shannon information [16,14]. It is important to stress that eq. (8) is a microscopic definition of entropy which coincides with the standard thermodynamic entropy only after maximization, see eq. (13) below.

If the system is characterized by L observables (or "extensive" variables[2]), $\hat{\mathbf{A}} = \{\hat{A}_\ell\}$, known in average $\langle\hat{A}_\ell\rangle = \mathrm{Tr}\,\hat{D}\hat{A}_\ell$, the variation is not free and one should maximize the constrained entropy

$$S' = S - \sum_\ell \lambda_\ell \langle\hat{A}_\ell\rangle, \tag{9}$$

where the $\boldsymbol{\lambda} = \{\lambda_\ell\}$ are L Lagrange multipliers associated with the L constraints $\langle\hat{A}_\ell\rangle$.

A maximization of the entropy under constraints gives a prediction for the minimum-biased density matrix (or "event distribution") which can be viewed as a generalization of Gibbs equilibrium:

$$\hat{D}_{\boldsymbol{\lambda}} = \frac{1}{Z_{\boldsymbol{\lambda}}} \exp -\boldsymbol{\lambda}\cdot\hat{\mathbf{A}}, \tag{10}$$

where $\boldsymbol{\lambda}\cdot\hat{\mathbf{A}} = \sum_{\ell=1}^{L}\lambda_\ell\hat{A}_\ell$ and where $Z_{\boldsymbol{\lambda}}$ is the associated partition sum insuring the normalization of $\hat{D}_{\boldsymbol{\lambda}}$:

$$Z_{\boldsymbol{\lambda}} = \mathrm{Tr}\exp -\boldsymbol{\lambda}\cdot\hat{\mathbf{A}}. \tag{11}$$

Using this definition, we can compute the associated equations of state (EoS):

$$\langle\hat{A}_\ell\rangle = \partial_{\lambda_\ell}\log Z_{\boldsymbol{\lambda}}. \tag{12}$$

The entropy associated with $\hat{D}_{\boldsymbol{\lambda}}$ is

$$S[\hat{D}_{\boldsymbol{\lambda}}] = \log Z_{\boldsymbol{\lambda}} + \sum_\ell \lambda_\ell\langle\hat{A}_\ell\rangle, \tag{13}$$

which has the structure of a Legendre transform between the entropy and the thermodynamic potential. To interpret the Gibbs ensemble as resulting from the contact with a reservoir or to guarantee the stationarity of eq. (10), it is often assumed that the observables $\hat{A}_\ell$ are conserved quantities such as the energy $\hat{H}$, the particle (or charge) numbers $\hat{N}_i$ or the angular momentum $\hat{L}$ [19]. However, there is no formal reason to limit the state variables to constants of motion. Even more, the introduction of not conserved quantities might be a way to take into account some non-ergodic aspects. Indeed, an additional constraint reduces the entropy, limiting the populated phase space or modifying the event distribution. This point will be developed at length in the next sections.

---

[1] In this article we implicitly use units such that the Boltzmann constant $k = 1$.

[2] In this paper the word "extensive" is used in the general sense of resulting from an observation, *i.e.* the $\langle\hat{A}_\ell\rangle$, and not in the restricted sense of additive variable. Intensive variables are conjugate to extensive variables *i.e.* Lagrange multipliers λ_ℓ imposing the average value of the associated extensive variable.

It should be noticed that microcanonical thermodynamics also corresponds to a maximization of the entropy (8) in a fixed energy subspace. In this case the maximum of the Shannon entropy can be identified with the Boltzmann entropy

$$\max\left(S\right)=\log W\left(E\right), \tag{14}$$

where W is the total state density with the energy E. The microcanonical case can also be seen as a particular Gibbs equilibrium (10) for which both the energy and its fluctuation are constrained. This so-called Gaussian ensemble in fact interpolates between the canonical and microcanonical ensembles depending upon the constraint on the energy fluctuation [24,51], and the same procedure can be applied to any conservation law. In this sense the Gibbs formulation (10) can be considered as the most general.

The ensemble of extensive variables constrained exactly or in average completely defines the statistical ensemble. This means that many different ensembles can be defined, and the most appropriate description of a finite system may be different from the standard microcanonical, canonical or grand-canonical.

4 Finite size and boundary conditions

An important problem when considering finite-size systems is the need to define boundary conditions to define the finite size. This is only a mathematical detail for "condensed" systems, *i.e.* finite-size self-bound systems in a much larger container, or particles trapped in an external confining potential [52]. In the other cases, finite-size systems can only be defined when proper boundary conditions are specified. Conversely to the thermodynamic limit which, when it exists, clearly isolates bulk properties independent of the actual shape of the container, finite-size systems explicitly depend on boundary conditions.

From a mathematical point of view the system Hamiltonian $\hat{H}$ is not defined until boundary conditions are specified. For example, for a particle problem a boundary can be the definition of a surface given by the implicit equation $\sigma(x,y,z)=0$. Since the Hamiltonian $\hat{H}_\sigma$ explicitly contains the boundary, the entropy S_σ also directly depends upon the definition of this boundary, according to

$$S_\sigma(E)=\log\operatorname{tr}\delta(E-\hat{H}_\sigma). \tag{15}$$

This brings a severe conceptual problem; the knowledge of the boundary requires an infinite information: the values of the function σ defining the actual surface in each space point. This is easily seen introducing the projector $\hat{P}_\sigma$ over the surface and its exterior. Indeed the boundary conditions applied to each microstate $\hat{P}_\sigma|\Psi^{(n)}\rangle=0$ is exactly equivalent to the extra constraint $\langle\hat{P}_\sigma\rangle=\operatorname{Tr}\hat{D}\hat{P}_\sigma=0$. If we note again $\hat{\mathbf{A}}$, the observables (including the Hamiltonian $\hat{H}_\sigma$) characterizing a given equilibrium, the density matrix including the boundary condition reads

$$\hat{D}_{\boldsymbol{\lambda}\sigma}=\frac{1}{Z_{\boldsymbol{\lambda}\sigma}}\exp-\boldsymbol{\lambda}\cdot\hat{\mathbf{A}}-b\hat{P}_\sigma\,, \tag{16}$$

which shows that the thermodynamics of the system does not only depend on the Lagrange multiplier b, but on the whole surface. For the very same global features such as the same average particle density or energy, we will have as many different thermodynamics as boundary conditions. More important, to specify the density matrix, the projector $\hat{P}_\sigma$ has to be exactly known and this is in fact impossible. The nature of $\hat{P}_\sigma$ is intrinsically different from the usual global observables $\hat{A}_\ell$. At variance with $\hat{A}_\ell$, $\hat{P}_\sigma$ is a many-body operator which does not correspond to any physical measurable observable. The knowledge of $\hat{P}_\sigma$ requires the exact knowledge of each point of the boundary surface while no or few parameters are sufficient to define $\hat{A}_\ell$. If we consider statistical physics as founded by the concept of minimum information [14,16], it is difficult to justify such an exact knowledge of the boundary. One should rather apply the minimum information concept also to the boundaries, introducing a hierarchy of collective observables which define the size and shape of the considered system. This amounts to introduce statistical ensembles treating the boundaries as additional extensive variables fixed by conjugated Lagrange parameters [53]. If, for instance, we consider that the relevant size information for an unbound system is its global square radius $\langle\hat{R}^2\rangle$, the adequate partition sum is

$$Z_\lambda=\sum_R W\left(R\right)e^{-\lambda R^2}, \tag{17}$$

where $W(R)$ is the state density associated with each possible value of the system radius.

5 Concept of equilibrium

As we have discussed in the previous sections, a statistical treatment is justified whenever a very large number of microstates exists for a given set of observables. This is always the case for the output of a heavy-ion collision, meaning that at least in principle a statistical approach should always be successful. An ensemble of events coming from similary prepared initial systems and/or selected by sorting always constitutes a statistical ensemble. Indeed, using the entropy concept, different observations are associated with a different information content. If we are able to recognize all the relevant degrees of freedom (*i.e.* the observations with a strong information content) the ensemble of replicas is by construction a statistical ensemble, *i.e.* a Gibbs equilibrium in the extended sense of sect. 3.

This generic statement hides the fundamental problem of recognizing the relevant observables. In the statistical models used to describe nuclear multifragmentation [54, 19] the hypothesis is made that all the information content is exhausted by the total energy, number of protons, neutrons, volume occupied at the time when fragments are decoupled, and in some cases angular momentum. This simple set of observables is certainly not sufficient to describe the whole phenomenology of heavy-ion reactions at all impact parameters. Then the remark is often made in the literature that "dynamical effects" dominate [55], meaning that extra constraints have to be put in order to have a statistical description of the final state of the reaction. However, the minimal information of multifragmentation statistical models [54,19] may be enough to describe limited portions of the collision phase space ("sources") properly selected by sorting [56]. The theoretical justification of this minimal statistical picture comes from the fact that complex classical systems subject to a non-linear dynamics are generally mixing [57]. In such a case the statistical ensemble is created by the propagation in time of initial fluctuations. The averages are averages over the initial conditions and the mixing character of the dynamics (if it can be proved) insures that the initial fluctuations are amplified in such a way that the ensemble of events covers the whole phase space uniformly[3]. For a classical dynamics which conserves the phase space volume of the ensemble of events, this means that the initial distribution is elongated and folded in such away that it gets close to any point of the phase space (the so-called baker transformation). This classical picture can be replaced in the quantum case by the idea of projection of irrelevant correlations [50]. The phase space can be described as a subspace of all possible observations. The regular quantum dynamics in the full space is transformed into a complex dynamics by the projection in the relevant observation subspace. Then two different realizations corresponding to the same projection, *i.e.* the same point in the relevant space, may differ in the full space (and consequently in their successive evolution) because of the unobserved correlations. This ensemble of correlations may lead to a statistical ensemble of realizations after a finite time. This phenomenon is often described introducing stochastic dynamics, *i.e.* assuming that the unobserved part of the dynamics which is averaged over is a random process [58,59].

5.1 How far is the system from equilibrium

An important point to be discussed is the justification of the statistical description. As we have just mentioned, the applicability of a statistical picture is in most cases an hypothesis (or a principle like in the thermodynamics second law). Therefore, the equilibrium hypothesis should be *a posteriori* controlled. Different properties can provide tests of equilibration such as

- the comparison with statistical models,
- the consistency of thermodynamical quantities, namely the compatibility of the different intensive variables measurements (*e.g.*, of the different thermometers) or the fulfillment of thermodynamic relations between averages and fluctuations (*e.g.*, $\sigma^2_{A_\ell} = \partial^2 \log Z_{\boldsymbol{\lambda}} / \partial\lambda_\ell^2 = -\partial\langle A_\ell\rangle/\partial\lambda_\ell$),
- the memory loss or the independence of the results on the preparation method of the considered ensemble.

However, it should be stressed that the real question is not whether the system is *at* equilibrium, but rather how far it is from a given equilibrium. Indeed, equilibrium is not unique in a finite system, and moreover exact equilibrium is a theoretical abstraction which cannot be achieved in the real world. To answer this question we should define a distance. The first idea could be to use the Liouville metric

$$d_{eq}^2 = \mathrm{tr}\left(\hat{D} - \hat{D}_{\boldsymbol{\lambda}}\right)^2 \tag{18}$$

between the actual ensemble characterized by the density matrix $\hat{D}$, and the equilibrium one $\hat{D}_{\boldsymbol{\lambda}}$ computed for the same collective variables $\langle A_\ell\rangle$. This is a nice theoretical tool, but a rather difficult definition as far as experiments are concerned. Another possibility is to introduce entropy as a metric [58]

$$d_{eq} = \left| S[\hat{D}] - S[\hat{D}_{\boldsymbol{\lambda}}] \right| / S[\hat{D}_{\boldsymbol{\lambda}}]. \tag{19}$$

This is a way to measure how far the system is from the maximum entropy state or in other words to measure how much information on the actual system is included in the collective variables $\{\langle A_\ell\rangle\}$ and how much is out of the considered equilibrium. This is a more physical distance but again it is difficult to implement in real experimental situations. A more practical measurement of the distance to equilibrium is to focus on the information used to deduce physical properties. Since the information about the actual system is contained in the observations $\langle\hat{O}_i\rangle$, the natural space to introduce this distance is the observation space. This is a formally well-defined problem since considering $\mathrm{Tr}\, \hat{O}_i\hat{O}_j$ as the scalar product between observables, the observation space has a well-defined topology. Then, when orthogonal observables are considered[4], the distance to equilibrium is simply

$$d_i = \left| \langle\hat{O}_i\rangle - \langle\hat{O}_i\rangle_{eq} \right|. \tag{20}$$

A typical example is given by the difference between the measured fluctuations $\sigma^2_{A_\ell} = \langle A_\ell^2\rangle - \langle A_\ell\rangle^2$ and the expected ones $\sigma^2_{A_\ell} = -\partial\langle A_\ell\rangle/\partial\lambda_\ell$ in the ensemble controlled by the λ_ℓ.

6 Finite systems and ensemble inequivalence

We have discussed that many different statistical ensembles can be defined when one considers finite systems. A fundamental theorem in statistical mechanics, the Van Hove theorem [60] (see appendix), guarantees the equivalence between different statistical ensembles at the thermodynamic limit. However the theorem does not apply in finite systems. In fact, it is strongly violated in first-order phase transitions if the system is finite, and this violation can persist up to the thermodynamic limit in the case of

---

[3] Meaning that any phase space point gets close to at least one event.

[4] If observables are not orthogonal it is alway possible to use a Schmitt procedure to define a set of orthogonal observables [16].

long-range forces. A consequence of that is that it is possible to give a rigorous definition of phase transitions even in finite systems, with the prediction of fancy phenomena like negative heat capacities, negative compressibilities and negative susceptibilities. The non-equivalence of statistical ensembles has also important conceptual consequences. It implies that the value of thermodynamic variables for the very same system depends on the type of experiment which is performed (*i.e.* on the ensemble of constraints which are put on the system), contrary to the standard thermodynamic viewpoint that water heated in a kettle is the same as water put in an oven at the same temperature. Ensemble inequivalence is the subject of an abundant literature (see, for example, refs. [22,23,25–30] for a discussion in a general context, and refs. [31–36] concerning phase transitions).

Generally speaking, for a given value of the control parameters (or *intensive variables*) λ_ℓ, the properties of a substance are univocally defined, *i.e.* the conjugated *extensive variables* $\langle \hat{A}_\ell \rangle$ have a unique value unambiguously defined by the corresponding equation of state ($\langle A_\ell \rangle = -\partial_{\lambda_\ell} \log Z(\{\lambda_\ell\})$). In reality, this fixes only the average value and the event-by-event value of the observation of $\hat{A}_\ell$ produces a probability distribution. The intuitive expectation that extensive variables at equilibrium have a unique value therefore means that the probability distribution is narrow and normal, such that a good approximation can be obtained by replacing the distribution with its most probable value. The normality of probability distributions is usually assumed on the basis of the central limit theorem. However, in finite systems the probability distributions has a finite width and moreover it can depart from a normal distribution. We will discuss in particular the case of a bimodal distribution [43]: in this case two different properties (phases) coexist for the same value of the intensive control variable.

The topological anomalies of probability distributions and the failure of the central limit theorem in phase coexistence imply that in a first-order phase transition the different statistical ensembles are in general not equivalent and different phenomena can be observed depending on the fact that the controlled variable is extensive or intensive. In the following, we will often take as a paradigm of intensive ensembles the canonical ensemble for which the inverse of the temperature β^{-1} is controlled, while the archetype of the extensive ensemble will be the microcanonical one for which energy is strictly controlled.

6.1 The difference between Laplace and Legendre

The relation between the canonical entropy and the logarithm of the partition sum is given by a Legendre transform eq. (13). It is important to distinguish between transformations within the same ensemble, as the Legendre transform, and transformations between different ensembles, which are given by non-linear integral transforms [35]. Let us consider energy as the extensive observable and inverse temperature β as the conjugated intensive one. The definition of the canonical partition sum is

$$Z_\beta = \sum_n \exp(-\beta E^{(n)}), \tag{21}$$

where the sum runs over the available eigenstates n of the Hamiltonian. Here, we assume that the partition sum converges; this is not always the case as discussed in ref. [61]. The possible divergence of the thermodynamic potential of the intensive ensemble is already a known case of ensemble inequivalence [19,61]. Computing the canonical (Shannon) entropy, we get

$$S_{can}(\langle E \rangle) = \log Z_\beta + \beta \langle E \rangle, \tag{22}$$

which is an exact Legendre transform since the EoS reads $\langle E \rangle = -\partial_\beta \log Z_\beta$. If energy can be treated as a continuous variable, eq. (21) can be written as

$$Z_\beta = \int_0^\infty \mathrm{d}E\ W(E) \exp(-\beta E), \tag{23}$$

where energies are evaluated from the ground state. Equation (23) is a Laplace transform between the canonical partition sum and the microcanonical density of states linked to the entropy by $S_E = \ln W(E)$. If the integrand $f(E) = \exp(E_E - \beta E)$ is a strongly peaked function, it can be approximated by a Gaussian (saddle point approximation) so that the integral can be replaced by the maximum $f(\bar{E})$ times a Gaussian integral. Neglecting this factor, we get

$$Z_\beta \approx W(\bar{E}) \exp(-\beta \bar{E}), \tag{24}$$

which can be rewritten as

$$\ln Z_\beta \approx S_{\bar{E}} - \beta \bar{E}; \tag{25}$$

or introducing the free energy $F_T = -\beta^{-1} \ln Z_\beta$,

$$F_T \approx \bar{E} - T S_{\bar{E}}. \tag{26}$$

Equation (25) has the structure of an approximate Legendre transform similar to the exact expression (22). This shows that in the lowest-order saddle point approximation eq. (24), the ensembles differing at the level of constraints acting on a specific observable (here energy) lead to the same entropy, *i.e.* they are equivalent. We will see in the next section that, however, the saddle point approximation eq. (24) can be highly incorrect close to a phase transition for the simple reason that the integrand is bimodal making a unique saddle point approximation inadequate. In this case eq. (25) cannot be applied, eq. (23) is the only correct transformation between the different ensembles, and ensemble inequivalence naturally arises.

6.2 Ensemble inequivalence and phase transitions

Let us consider the case of a first-order phase transition where the canonical energy distribution

$$P_{\beta_0}(E) = W(E) \exp(-\beta_0 E)/Z_{\beta_0} \tag{27}$$

has a characteristic bimodal shape [43,46,47] at the temperature β_0 with two maxima $\overline{E}_\beta^{(1)}$, $\overline{E}_\beta^{(2)}$ that can be associated with the two phases. It is easy to see that eq. (23) can also be seen as a Laplace transform of the canonical probability $P_{\beta_0}(E)$

$$Z_\beta = Z_{\beta_0} \int_0^\infty \mathrm{d}E \; P_{\beta_0}(E) \exp(-(\beta - \beta_0)E). \tag{28}$$

A single saddle point approximation is not valid when $P_{\beta_0}(E)$ is bimodal; however it is always possible to write

$$P_\beta = m_\beta^{(1)} P_\beta^{(1)} + m_\beta^{(2)} P_\beta^{(2)}, \tag{29}$$

with $P_\beta^{(i)}$ mono-modal normalized probability distribution peaked at $\overline{E}_\beta^{(i)}$. The canonical mean energy is then the weighted average of the two energies

$$\langle E \rangle_\beta = \tilde{m}_\beta^{(1)} \overline{E}_\beta^{(1)} + \tilde{m}_\beta^{(2)} \overline{E}_\beta^{(2)}, \tag{30}$$

with

$$\tilde{m}_\beta^{(i)} = m_\beta^{(i)} \int \mathrm{d}E P_\beta^{(i)}(E) E / \overline{E}_\beta^{(i)} \simeq m_\beta^{(i)}. \tag{31}$$

Since only one mean energy is associated with a given temperature β^{-1}, the canonical caloric curve is monotonic, and the microcanonical one is not. Indeed it is immediate to see from eq. (27) that the bimodality of P_β implies then a back bending of the microcanonical caloric curve $T^{-1} = \partial_E S$, meaning that in the first-order phase transition region the two ensembles are not equivalent. If instead of looking at the average $\langle E \rangle_\beta$ we look at the most probable energy $\overline{E}_\beta$, this (unusual) canonical caloric curve is identical to the microcanonical one, up to the transition temperature β_t^{-1} for which the two components of $P_\beta(E)$ have the same height. At this point the most probable energy jumps from the low- to the high-energy branch of the microcanonical caloric curve.

The question arises whether this violation of ensemble equivalence survives towards the thermodynamic limit. This limit can be expressed as the fact that the thermodynamic potentials per particle converge when the number of particles N goes to infinity:

$$f_{N,\beta} = \beta^{-1} \frac{\log Z_\beta}{N} \to \bar{f}_\beta; \quad s_N(e) = \frac{S(E)}{N} \to \bar{s}(e), \tag{32}$$

where $e = E/N$. Let us also introduce the reduced probability $p_{N,\beta}(e) = (P_\beta(N,E))^{1/N}$ which then converges towards an asymptotic distribution

$$p_{N,\beta}(e) \to \bar{p}_\beta(e); \quad \bar{p}_\beta(e) = \exp\left(\bar{s}(e) - \beta e + \bar{f}_\beta\right). \tag{33}$$

Since $P_\beta(N,E) \approx (\bar{p}_\beta(e))^N$, one can see that when $\bar{p}_\beta(e)$ is normal, the relative energy fluctuation in $P_\beta(N,E)$ is suppressed by a factor $1/\sqrt{N}$. At the thermodynamic limit P_β reduces to a δ-function and ensemble equivalence is recovered. To analyze the thermodynamic limit of a bimodal $p_{N,\beta}(e)$, let us introduce as before $\beta_{N,t}^{-1}$ the temperature for which the two maxima of $p_{N,\beta}(e)$ have the same height. For a first-order phase transition $\beta_{N,t}^{-1}$ converges to a fixed point $\bar{\beta}_t^{-1}$ as well as the two maximum energies $e_{N,\beta}^{(i)} \to \bar{e}_\beta^{(i)}$. For all temperatures lower (higher) than $\bar{\beta}_t^{-1}$ only the low- (high-) energy peak will survive at the thermodynamic limit, since the difference of the two maximum probabilities will be raised to the power N. Therefore, below $\bar{e}_\beta^{(1)}$ and above $\bar{e}_\beta^{(2)}$ the canonical caloric curve coincides with the microcanonical one in the thermodynamic limit. In the canonical ensemble the temperature $\bar{\beta}_t^{-1}$ corresponds to a discontinuity in the state energy irrespectively of the behavior of the entropy between $\bar{e}_\beta^{(1)}$ and $\bar{e}_\beta^{(2)}$.

The microcanonical caloric curve in the phase transition region may either converge towards the Maxwell construction, or keep a backbending behavior [21], since a negative heat capacity system can be thermodynamically stable even in the thermodynamic limit if it is isolated [25]. Examples of a backbending behavior at the thermodynamic limit have been reported for a model many-body interaction taken as a functional of the hypergeometric radius in the analytical work of ref. [3], and for the long-range Ising model [4]. This can be understood as a general effect of long-range interactions for which the topological anomaly leading to the convex intruder in the entropy is not cured by increasing the number of particles [4,62]. Conversely, for short-range interactions [15] the backbending is a surface effect which should disappear at the thermodynamic limit. This is the case for the Potts model [32], the microcanonical model of fragmentation of atomic clusters [63] and for the lattice gas model with fluctuating volume [48]. The interphase surface entropy goes to zero as $N \to \infty$ in these models, leading to a linear increase of the entropy in agreement with the canonical predictions.

Within the approach based on the topology of the probability distribution of observables [43] it was shown that ensemble inequivalence arises from fluctuations of the order parameter [22]. Ensembles putting different constraints on the fluctuations of the order parameter lead to a different thermodynamics. In the case of phase transitions with a finite latent heat, the total energy usually plays the role of an order parameter except in the microcanonical ensemble which, therefore, is expected to present a different thermodynamics than the other ensembles [19]. This inequivalence may remain at the thermodynamic limit if the involved phenomena are not reduced to short-range effects.

6.3 Temperature jump at constant energy

In particular, it may happen that the energy of a subsystem becomes an order parameter when the total energy is constrained by a conservation law or a microcanonical sorting. This frequently occurs for Hamiltonians containing a kinetic energy contribution [3,4,64]: if the kinetic heat capacity is large enough, it becomes an order parameter in the microcanonical ensemble. Then, the microcanonical caloric curve presents at the thermodynamical limit a temperature jump in complete disagreement with the canonical ensemble.

To understand this phenomenon, let us consider a finite system for which the Hamiltonian can be separated into two components $E = E_1 + E_2$, that are statistically independent $(W(E_1, E_2) = W_1(E_1)W_2(E_2))$ and such that the associated degrees of freedom scale in the same way with the number of particles; we will also consider the case where $S_1 = \log W_1$ has no anomaly while $S_2 = \log W_2$ presents a convex intruder [15] which is preserved at the thermodynamic limit. Typical examples of E_1 are given by the kinetic energy for a classical system with velocity-independent interactions, or other similar one-body operators [4]. The probability to get a partial energy E_1 when the total energy is E is given by

$$P_E(E_1) = \exp\left(S_1(E_1) + S_2(E - E_1) - S(E)\right). \quad (34)$$

The extremum of $P_E(E_1)$ is obtained for the partitioning of the total energy E between the kinetic and potential components that equalizes the two partial temperatures

$$\overline{T_1}^{-1} = \partial_{E_1} S_1(\overline{E}_1) = \partial_{E_2} S_2(E - \overline{E}_1) = \overline{T_2}^{-1}. \quad (35)$$

If $\overline{E}_1$is unique, $P_E(E_1)$ is mono-modal and we can use a saddle point approximation around this solution to compute the entropy

$$S(E) = \log \int_{-\infty}^{E} \mathrm{d}E_1 \exp\left(S_1(E_1) + S_2(E - E_1)\right). \quad (36)$$

At the lowest order, the entropy is simply additive so that the microcanonical temperature of the global system $\partial_E S(E) = \overline{T}^{-1}$ is the one of the most probable energy partition. Therefore, the most probable partial energy $\overline{E}_1$ acts as a microcanonical thermometer. If $\overline{E}_1$ is always unique, the kinetic thermometer in the backbending region will follow the whole decrease of temperature as the total energy increases. Therefore, the total caloric curve will present the same anomaly as the potential one. If, conversely, the partial energy distribution is double humped [65], then the equality of the partial temperatures admits three solutions, one of them $\overline{E}_1^{(0)}$ being a minimum. At this point, the partial heat capacities

$$C_1^{-1} = -\overline{T}^2\, \partial^2_{E_1} S_1(\overline{E}_1^{(0)}); \quad C_2^{-1} = -\overline{T}^2\, \partial^2_{E_2} S_2(E - \overline{E}_1^{(0)}) \quad (37)$$

fulfill the relation

$$C_1^{-1} + C_2^{-1} < 0. \quad (38)$$

This happens when the potential heat capacity is negative and the kinetic energy is large enough $(C_1 > -C_2)$ to act as an approximate heat bath: the partial energy distribution $P_E(E_1)$ in the microcanonical ensemble is then bimodal as the total energy distribution $P_\beta(E)$ in the canonical ensemble, implying that the kinetic energy is the order parameter of the transition in the microcanonical ensemble. In this case the microcanonical temperature is given by a weighted average of the two estimations from the two maxima of the kinetic energy distribution

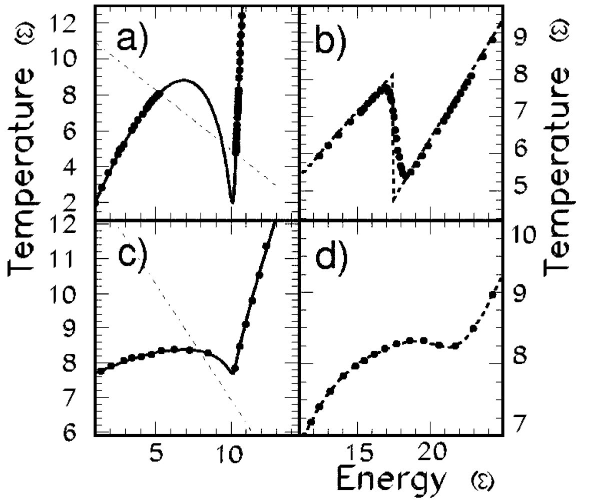

Fig. 2. Left panels: temperature as a function of the potential energy E_2 (full lines) and of the kinetic energy $E - E_2$ (dot-dashed lines) for two model equation of states of classical systems showing a first-order phase transition. Symbols: temperatures extracted from the most probable kinetic energy thermometer from eq. (35). Right panels: total caloric curves (symbols) corresponding to the left panels and thermodynamic limit of eq. (39) (dashed lines).

$$T = \partial_E S(E) = \frac{\overline{P}^{(1)}\sigma^{(1)}/\overline{T}^{(1)} + \overline{P}^{(2)}\sigma^{(2)}/\overline{T}^{(2)}}{\overline{P}^{(1)}\sigma^{(1)} + \overline{P}^{(2)}\sigma^{(2)}}, \quad (39)$$

where $\overline{T}^{(i)} = T_1(\overline{E}_1^{(i)})$ are the kinetic temperatures calculated at the two maxima, $\overline{P}^{(i)} = P_E(\overline{E}_1^{(i)})$ are the probabilities of the two peaks and $\sigma^{(i)}$ their widths. At the thermodynamic limit eq. (38) reads $c_1^{-1} + c_2^{-1} < 0$, with $c = \lim_{N\to\infty} C/N$. If this condition is fulfilled, the probability distribution $P_\beta(E)$ presents two maxima for all finite sizes and only the highest peak survives at $N = \infty$. Let E_t be the energy at which $P_{E_t}(\overline{E}^{(1)}) = P_{E_t}(\overline{E}^{(2)})$. Because of eq. (39), at the thermodynamic limit the caloric curve will follow the high- (low-) energy maximum of $P_E(E_1)$ for all energies below (above) E_t; there will be a temperature jump at the transition energy E_t.

Let us illustrate the above results with two examples for a classical gas of interacting particles. For the kinetic energy contribution we have $S_1(E) = c_1 \ln(E/N)^N$ with a constant kinetic heat capacity per particle $c_1 = 3/2$. For the potential part we will take two polynomial parametrizations of the interaction caloric curve presenting a back bending which are displayed in the left part of fig. 2 in units of an arbitrary scale ϵ. If the decrease of the partial temperature $T_2(E_2)$ is steeper than $-2/3$ (fig. 2a) then eq. (38) is verified [3] and the kinetic caloric curve $T_1(E - E_1)$ (dot-dashed line) crosses the potential one $T_2(E_2)$ (full line) in three different points for all values

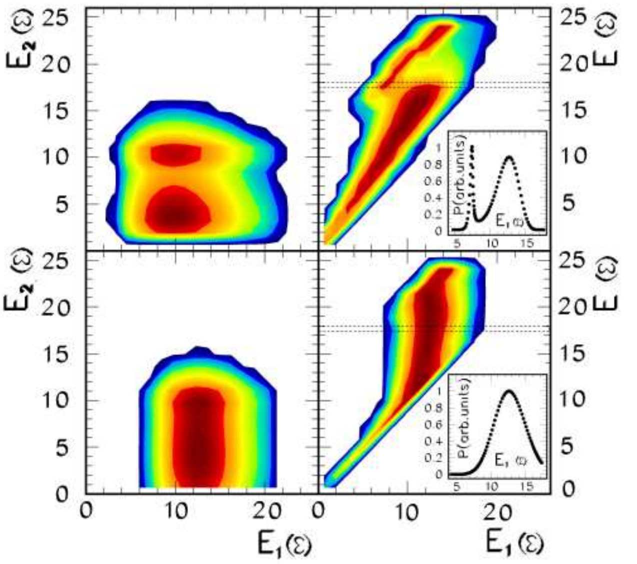

Fig. 3. Canonical event distributions in the potential *versus* kinetic energy plane (left panels) and total *versus* kinetic energy plane (right panels) at the transition temperature for the two model equations of state of fig. 2. The inserts show two constant total energy cuts of the distributions.

of the total energy lying inside the region of coexistence of two kinetic energy maxima. The resulting caloric curve for the whole system is shown in fig. 2b (symbols) together with the thermodynamic limit (line) evaluated from the double saddle point approximation eq. (39). In this case one observes a temperature jump at the transition energy. If the temperature decrease is smoother (fig. 2c) the shape of the interaction caloric curve is preserved at the thermodynamic limit (fig. 2d).

The occurrence of a temperature jump in the thermodynamic limit is easily spotted by looking at the bidimensional canonical event distribution $P_\beta(E_1, E_2)$ shown at the transition temperature $\beta = \beta_t$ in the left part of fig. 3 for the two model equation of states of fig. 2. In the canonical ensemble the kinetic energy distribution is normal. These same distributions are shown as a function of E and E_1, $P_\beta(E, E_1) \propto \exp S_1(E_1) \exp S_2(E - E_1) \exp(-\beta E)$ in the right part of fig. 3. The microcanonical ensemble is a constant energy cut of $P_\beta(E, E_1)$, which leads to the microcanonical distribution $P_E(E_1)$ within a normalization constant. If the anomaly in the potential equation of state is sufficiently important, the distortion of events due to the coordinate change is such that one can still see the two phases coexist even after a sorting in energy.

7 Definitions of phase transitions

Phase transitions are universal properties of matter in interaction. In macroscopic physics, they are singularities (*i.e.* non-analytical behaviors) in the system equation of state (EoS) and hence classified according to the degree of non-analyticity of the EoS at the transition point. Then, a phase transition is an intrinsic property of the system and not of the statistical ensemble used to describe the equilibrium. Indeed, at the thermodynamic limit all the possible statistical ensembles converge towards the same EoS (see appendix), and the various thermodynamic potentials are related by simple Legendre transformations leading to a unique thermodynamics. On the other side for finite systems, as discussed above, two ensembles which put different constraints on the fluctuations of the order parameter lead to qualitatively different EoS close to a first-order phase transition [15,24]. Thermodynamic observables like heat capacities can, therefore, be completely different depending on the experimental conditions of the measurement. Moreover, such inequivalences may survive at the thermodynamic limit if forces are long ranged as for self-gravitating objects [3,4]. In fact, the characteristic of phase transitions in finite systems, and in particular the occurrence of a negative heat capacity, have first been discussed in the astrophysical context [2,25,30,66–70]. Since these pioneering works in astrophysics, an abundant literature is focused on the understanding of phase transitions in small systems from a general point of view [15,29,35, 71–76] or in the mean-field context [4,77] or for some specific systems such as metallic clusters [47,65] or nuclei [78] and even DNA [79].

7.1 Phase transitions in infinite systems

Let us first recall the definition of phase transitions in infinite systems. At the thermodynamic limit for short-range interactions the statistical ensembles are equivalent and it is enough to reduce the discussion to the ensemble where only one extensive variable A_L is kept fixed, all the others being constrained through the associated Lagrange parameters. The typical example is the grand-canonical ensemble where only the volume $A_L = V$ is kept as an extensive variable. Then all the thermodynamics is contained in the associated potential $\log Z_{\lambda_1,\ldots,\lambda_{L-1}}(A_L)$. Since it is extensive, the potential is proportional to the remaining extensive variable

$$\log Z_{\lambda_1,\ldots,\lambda_{L-1}}(A_L) = A_L \lambda_L(\lambda_1, \ldots, \lambda_{L-1}) \tag{40}$$

so that all the non-trivial thermodynamic properties are included in the reduced potential, *i.e.* the intensive variable

$$\lambda_L = \partial_{A_L} \log Z_{\lambda_1,\ldots,\lambda_{L-1}}(A_L) = \frac{\log Z_{\lambda_1,\ldots,\lambda_{L-1}}}{A_L} \tag{41}$$

associated with A_L. In the grand-canonical case $A_L = V$, the reduced potential is the pressure, $\lambda_L \propto P$, which is then a function of the temperature and the chemical potential(s). In this limit all the thermodynamics is included in the single function $\lambda_L(\lambda_1, \ldots, \lambda_{L-1})$, and this is why in the literature $p(V)$ is often loosely referred to as "the" EoS, and the existence of many EoS is ignored. If this EoS is analytical, all the thermodynamic quantities which are all derivatives of the thermodynamic potential, present smooth behaviors, and no phase transition appears. A phase transition is a major modification of the macrostate

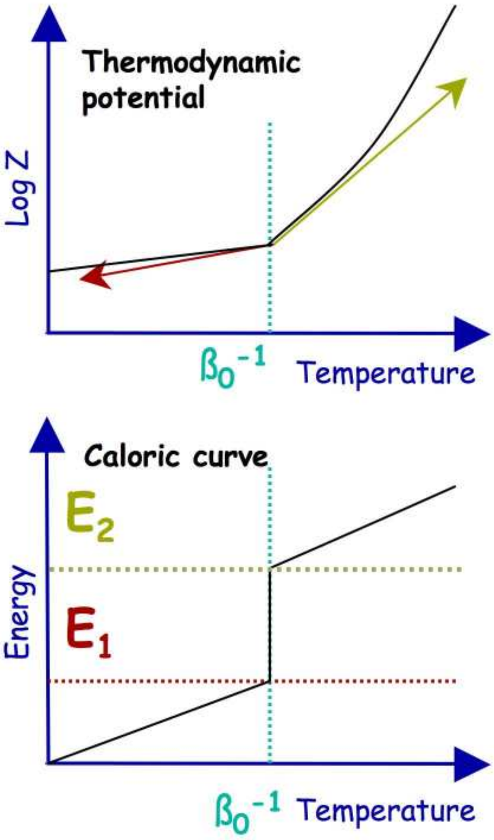

Fig. 4. Schematic representation of a first-order phase transition in the canonical case. Top: the log of the canonical partition sum (*i.e.* the free energy) presents an angular point. Bottom: the first derivative as a function of the temperature (*i.e.* the energy) presents a jump.

properties for a small modification of the control parameters $(\lambda_1, \ldots, \lambda_{L-1})$. Such an anomalous behavior can only happen if the thermodynamic potential presents a singularity. This singularity can be classified according to the order of the derivative which presents a discontinuity or a divergence. According to Ehrenfest this is the order of the phase transition. In modern statistical mechanics, all the higher-order transitions are called under the generic name of continuous transitions. Figure 4 schematically illustrates a first-order phase transition in the canonical ensemble.

7.2 Phase transitions in finite systems

As soon as one considers a finite physical system, all the above discussion does not apply. First the thermodynamic potential and observables are not additive, therefore we cannot introduce a reduced potential. Indeed, the inequality

$$\lambda_L(\lambda_1, \ldots, \lambda_{L-1}, A_L) \equiv \frac{\partial \log Z_{\lambda_1,\ldots,\lambda_{L-1}}(A_L)}{\partial A_L} \neq \frac{\log Z_{\lambda_1,\ldots,\lambda_{L-1}}(A_L)}{A_L} \qquad (42)$$

shows that the grand potential per unit volume does no longer give the pressure and presents a non-trivial volume dependence. Moreover, the analysis of the singularities of the thermodynamic potential has no meaning, since it is an analytical function. The standard statistical-physics textbooks thus conclude that rigourously speaking there is no phase transitions in finite systems. However, as we have already mentioned, first for self-gravitating objects [2,25, 30,66–70] and then in small systems [3,15,29,35,47,71–73, 76,78] it was shown that phase transitions might be associated with the occurrence of negative microcanonical heat capacities. This can be generalized to the occurence of an inverted curvature of the thermodynamic potential of any ensemble keeping at least one extensive variable A_L not orthogonal to the order parameter[5] [35,80]. In the following we call this ensemble an extensive ensemble. Then, negative compressibility or negative susceptibility should be, like negative heat capacity, observed in first-order phase transitions of finite systems. In the microcanonical ensemble of classical particles, it was proposed that anomalously large fluctuations of the kinetic energy, *i.e.* larger than the expected canonical value, highlight a negative heat capacity [81]. It was then demonstrated that those two signals of a phase transition, negative curvatures and anomalous fluctuations, observed in extensive ensembles where the order parameter is fixed, are directly related to the appearance of bimodalities in the distribution of this order parameter in the intensive ensemble where the order parameter is only fixed in average through its conjugated Lagrange multiplier [6,43].

The occurrence of bimodalities is discussed in the literature since a long time and is often used as a practical way to look for phase transitions in numerical simulations [17, 46]; however, the general equivalence between negative curvatures and bimodalities was presented in ref. [43]. For intensive ensembles, since the pioneering work of Yang and Lee [42] another definition was proposed considering the zeroes of the partition sum in the complex intensive parameter plane [42,82]. The idea is simple: the zeroes of Z are the singularities of $\log Z$ and so phase transitions, which are singularities, must come from the zeroes of the partition sum. In a finite system the zeroes of the partition sum cannot be on the real axis since the partition sum Z is the sum of exponential factors which cannot produce a singularity of $\log Z$. However, the thermodynamic limit of an infinite volume may bring the singularity on the real axis. This is schematically illustrated in fig. 5. Only regions where zeroes converge towards the real axis may present phase transitions, while the other regions present no anomalies. The order of the transition can be associated with the asymptotic behavior of the zeroes [82].

The distribution of zeroes has been analyzed in ref. [44] where the transition was studied with a parabolic entropy. In ref. [45] the equivalence of the expected behavior of the zeroes in a first-order phase transition case and the occurrence of bimodalities in the distribution of the associated extensive parameter was demonstrated. To be precise, in

---

[5] Orthogonality is here defined using the trace as a scalar product between observables following sect. 3.

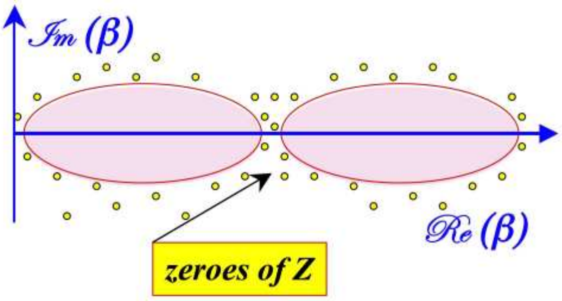

Fig. 5. Schematic representation of the zeroes of the partition sum Z in the complex temperature plane. The regions where no zeroes are coming close to the real axis when the thermodynamic limit is taken will not present singularities of $\log Z$.

this demonstration bimodality means that the extensive variable distribution can be split at the transition point into two distributions of equal height, with the distance between the two maxima scaling like the system size [83].

This global picture of phase transitions in finite systems is summarized in fig. 6 in the case where energy is the order parameter of the transition. The occurrence of a bimodal distribution of the extensive parameter (*e.g.*, energy) in the associated intensive (*e.g.*, canonical) ensemble is a necessary and sufficient condition to asymptotically get the distribution of the Yang-Lee zeroes in the complex Lagrange multiplier (*e.g.*, temperature) plane, which is expected in a first-order transition. The direction of bimodality is the direction of the order parameter. This bimodality is also equivalent to the presence of an anomalous curvature in the thermodynamic potential of the extensive (microcanonical) ensemble obtained constraining the bimodal observable to a fixed value. In the extensive ensemble, the inverted curvature can be spotted looking for anomalously large fluctuations (*e.g.*, larger than the canonical ones) of the partition of the extensive variable (*e.g.*, energy) between two independent subsystems.

8 Statistical description of evolving systems

A major issue in the statistical treatment of finite systems is that most of the time open and transient systems are studied. Therefore, they are not only finite in size but also finite in time and, in fact, they are evolving. The number of degrees of freedom of a quantum many-body problem being infinite, it is impossible to have all the information needed to solve exactly the dynamical problem. Since only a small part of the observation space is relevant, this time evolution may also be treated with statistical tools. This is the purpose of many models: from Langevin approaches to Fokker-Planck equations, from hydrodynamics to stochastic Time-Dependent Hartree-Fock theory. The purpose of this paper is not to review those theoretical approaches, therefore we will not enter here into details about the different recent progresses, and we will rather focus this discussion on general arguments of time-dependent statistical ensembles [50,53].

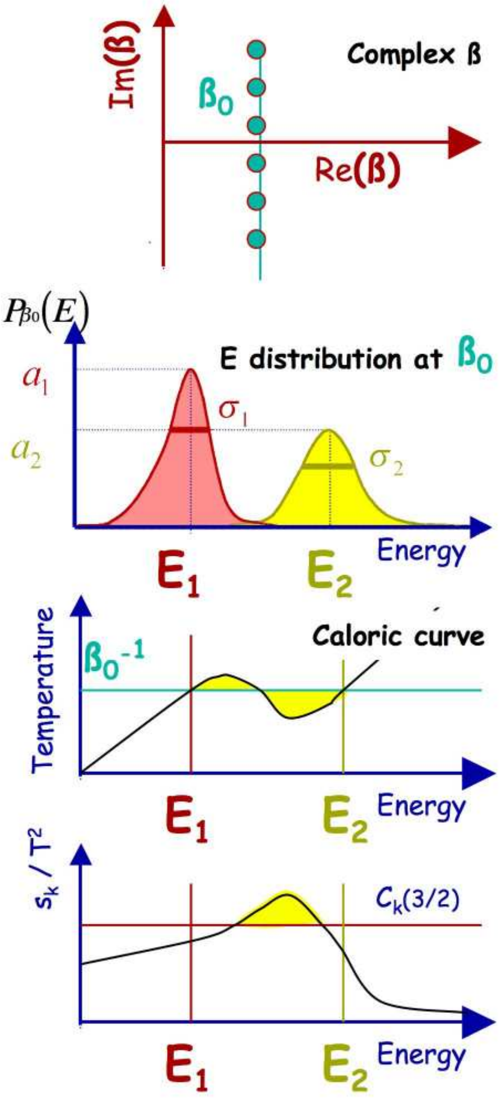

Fig. 6. Schematic representation of the different equivalent definitions of first-order phase transitions in finite systems. From top to bottom: the partition sum's zeroes aligning perpendicular to the real temperature axis with a density scaling like the number of particles; the bimodality of the energy distribution with a distance between the two maxima scaling like the number of particles times the latent heat; the appearence of a back-bending in the microcanonical caloric curve, *i.e.* a negative heat capacity region; and the observation of anomalously large fluctuations of the energy splitting between the kinetic part and the interaction part.

A statistical treatment of a dynamical process is based on the idea that at any time one can consider only the relevant variables A_ℓ, disregarding all the other ones a_m as irrelevant. If only the maximum entropy state is followed in time assuming that all the irrelevant degrees of freedom have relaxed instantaneously, one gets a generalized mean-field approach [58]. If the fluctuations of the irrelevant degrees of freedom are included, this leads to a Langevin dynamics [59]. With those considerations one can see that statistical approaches can be always improved including more and more degrees of freedom to asymptotically become exact. However, before including a huge number of degrees of freedom one should ask himself if only a few observables can take care of the most important dynamical aspects of the systems we are looking at. In a recent paper [53] it was proposed to introduce observations at different times (*e.g.*, different freeze-out/equilibration times) as well as time-odd extensive parameters. The idea is simple: maximizing the Shannon entropy with different observables $\hat{A}_\ell$ known at different times $t_\ell = t_0 + \Delta t_\ell$ is a way to treat a part of the dynamics. Going to the Heisenberg representation, if we propagate all the $\hat{A}_\ell$ to the same time t_0 we get

$$\hat{A}_\ell(t_0) = e^{-i\Delta t_\ell \hat{H}} \hat{A}_\ell e^{i\Delta t_\ell \hat{H}} \tag{43}$$

$$= \hat{A}_\ell - i\Delta t_\ell [\hat{H}, \hat{A}_\ell] + \dots . \tag{44}$$

This shows that the time propagation introduces new constraining operators

$$\hat{B}_\ell = -i[\hat{H}, \hat{A}_\ell]. \tag{45}$$

If $\hat{A}_\ell$ is a time-even observable, $\hat{B}_\ell$ is a time-odd operator. Let us take the example of an unconfined finite system characterized at a given time by a typical size $\langle \hat{R}^2 \rangle = \langle \hat{S} \rangle$, where $\hat{R}^2$ is the one body operator $\sum_i \hat{r}_i^2$. If the whole information is assumed to be known at the same time, then the statistical distribution of events reads in a classical canonical picture

$$p^{(n)} = \frac{1}{Z} e^{-\beta E^{(n)} - \lambda S^{(n)}}, \tag{46}$$

which is formally equivalent to a particle in a harmonic potential. However, if now we assume that the size information is coming from a different time, then according to eq. (45) we must introduce a new time-odd operator $\hat{v}_r = -i[\hat{H}, \hat{r}^2]$. For a local interaction, this reduces to

$$\hat{v}_r = (\hat{r}\hat{p} + \hat{p}\hat{r})/m\,, \tag{47}$$

which represents a radial flow. Then the classical canonical probability reads

$$p^{(n)} = \frac{1}{Z} e^{-\beta\left(\mathbf{p}^{(n)} - h(t)\mathbf{r}^{(\mathbf{n})}\right)^2 - \lambda S^{(n)}} \tag{48}$$

which is a statistical ensemble of particles under a Hubblian flow. In the ideal-gas model eq. (48) provides the exact solution at any time of the dynamics. This simple example shows that information theory allows to treat in a statistical picture dynamical processes where observables are defined at different times, by taking into account time-odd components such as flows. This might be a tool to extract thermodynamical quantities from complex dynamics. In particular, the above example shows that in an open system an initial extension in space is always transformed into an expansion, meaning that flow is an essential ingredient even in statistical approaches.

9 Conclusion

In conclusion, we have presented in this paper the actual understanding of the thermodynamics of finite systems from the point of view of information theory. We have put some emphasis on first-order phase transitions which are associated with specific and intriguing phenomena as bimodalities and negative heat capacities. Phase transitions have been widely studied in the thermodynamic limit of infinite systems. However, in the physical situations considered here, this limit cannot be taken and phase transitions should be reconsidered from a more general point of view. This is for example the case of matter under long-range forces like gravitation. Even if these self-gravitating systems are very large they cannot be considered as infinite because of the non-saturating nature of the force. Other cases are provided by microscopic or mesoscopic systems built out of matter which is known to present phase transitions. Metallic clusters can melt before being vaporized. Quantum fluids may undergo Bose condensation or a super-fluid phase transition. Dense hadronic matter should merge in a quark and gluon plasma phase while nuclei are expected to exhibit a liquid-gas phase transition and a superfluid phase. For all these systems the theoretical and experimental issue is how to define and sign a possible phase transition in a finite system. In this review we have presented the synthesis of different works which tend to show that phase transitions can be defined as clearly as in the thermodynamic limit. Depending upon the statistical ensemble, *i.e.* on the experimental situation, one should look for different signals. In the ensemble where the order parameter is free to fluctuate (intensive ensemble), the topology of the event distribution should be studied. A bimodal distribution signals a first-order phase transition. The direction in the observable space in which the distribution is bimodal defines the best order parameter. To survive the thermodynamic limit, the distance between the two distributions, the two "phases", should scale like the number of particles. This occurrence of a bimodal distribution is equivalent to the alignment of the partition sum zeroes as described by the Yang and Lee theorem. In the associated extensive ensemble, the bimodality condition is also equivalent to the requirement of a convexity anomaly in the thermodynamic potential. The first experimental evidences of such a phenomenon have been reported recently in different fields: the melting of sodium clusters [6], the fragmentation of hydrogen clusters [8], the pairing in nuclei [9] and nuclear multifragmentation [7, 84,85]. However, much more experimental and theoretical studies are now expected to progress in this new field of phase transitions in finite systems. Three challenges can thus be assigned to the physics community:

- The statistical description of non-extensive systems and in particular of open transient finite systems.
- The experimental and theoretical study of phase transitions in those systems and of the expected abnormal thermodynamics.
- The confirmation of the observation of the nuclear phase transition and the analysis of the associated

equation-of-state properties and the associated phase diagram.

Appendix A. The Van Hove theorem

Let us consider a system in a volume V for which only the average value of energy and number of particles is defined (grand-canonical ensemble). Let us calculate the grand potential $\Omega = -T \ln Z$:

$$Z_{\beta\mu}(V) = \sum_n \exp\left(-\beta\left(H^{(n)} - \mu N^{(n)}\right)\right), \qquad \text{(A.1)}$$

where the sum extends over all the possible configurations of the system, $H^{(n)} = K^{(n)} + U^{(n)}$ ($N^{(n)}$) represents the energy (number of particles) of the system in the configuration (n), and β, μ are the associated Lagrange multipliers, the inverse temperature and the chemical potential, respectively. The partition sum results

$$Z_{\beta\mu}(V) = \sum_{N=0}^{\infty} z_k^N Z_\beta(N,V) \qquad \text{(A.2)}$$

with $z_k = \exp(\beta\mu)(\frac{2m\pi}{h^2\beta})^{3/2}$ the ideal-gas part and

$$Z_\beta(N,V) = \frac{1}{N!}\int_V \mathrm{d}^{3N}r \exp\left(-\beta U\right) \qquad \text{(A.3)}$$

the partition sum associated with the interaction part. Let us divide $V = mV_0 + V_1$ in m equal boxes of volume V_0 separated by "corridors" of width b larger than the range of the force such that the interactions among particles in different boxes can be neglected (see fig. 7). The volume excluded by the corridors is V_1. To calculate $Z_\beta(N,V)$ let us consider the number of particles in the corridor N_1:

$$Z_\beta(N,V) = \sum_{N_1=0}^{N} \frac{1}{N_1!}\frac{1}{(N-N_1)!}\int_{V_1}\mathrm{d}^{3N_1}r\,, \qquad \text{(A.4)}$$

$$\int_{V-V_1}\mathrm{d}^{3(N-N_1)}r \exp\left(-\beta U\right). \qquad \text{(A.5)}$$

Let us note ϵ the minimum of the two-body interaction (see fig. 7); the potential energy in the corridor satisfies then the inequality $U_{V_1} \geq \varepsilon\xi N_1$, where $\xi = (b/a)^3$ represents the maximum number of particles interacting with a given particle. For the total potential energy, we can write

$$U \geq \varepsilon\xi N_1 + \frac{1}{2}\sum_{i=N_1+1}^{N} \qquad \text{(A.6)}$$

leading to

$$Z_\beta(N,V) \leq \sum_{N_1=0}^{N}\frac{1}{N_1!}\frac{1}{(N-N_1)!}V_1^{N_1}\exp\left(-N_1\beta\varepsilon\xi\right), \qquad \text{(A.7)}$$

$$\int_{mV_0}\mathrm{d}^{3(N-N_1)}r \exp\left(-\beta U\right), \qquad \text{(A.8)}$$

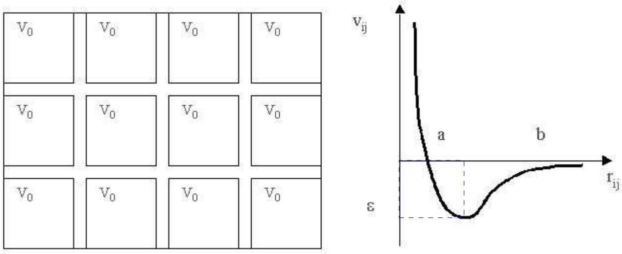

Fig. 7. Schematic representation of the Van Hove theorem demonstration (left) and the corresponding inter-particle interaction (right).

where the last integrals run over the n-independent volumes V_0. Introducing this expression in (A.4) with $N_2 = N - N_1$, the partition sum $Z_{\beta\mu}(V)$ reads

$$Z_{\beta\mu}(V) \leq \sum_{N_1=0}^{\infty}\frac{1}{N_1!}V_1^{N_1}z_k^{N_1}\exp\left(-N_1\beta\varepsilon\xi\right)$$

$$\cdot\sum_{N_2=0}^{\infty}\frac{1}{N_2!}z_k^{N_2}\int_{mV_0}\mathrm{d}^{3N_2}r\exp\left(-\beta U\right) \qquad \text{(A.9)}$$

$$= \exp\left(z_k V_1 e^{-\beta\varepsilon\xi}\right) Z_{\beta\mu}^m(V_0), \qquad \text{(A.10)}$$

where the last equality stems from the fact that particles interact only within the same box again because of the short range of the force. Finally, we get using $V_1 \propto mV_0^{2/3}$:

$$\log Z_{\beta\mu}(V) \leq kmV_0^{2/3} + m\log Z_{\beta\mu}(V_0), \qquad \text{(A.11)}$$

$$\frac{\log Z_{\beta\mu}(V)}{V} \leq kV_0^{-1/3} + \frac{\log Z_{\beta\mu}(V_0)}{V_0}, \qquad \text{(A.12)}$$

which gives in the thermodynamic limit (keeping m constant) $V \to \infty$, $V_0 \to \infty$, $V \to mV_0$,

$$\frac{\log Z_{\beta\mu}(V)}{V} \leq \frac{\log Z_{\beta\mu}(V_0)}{V_0}. \qquad \text{(A.13)}$$

On the other side, the opposite inequality is trivially true:

$$Z_{\beta\mu}(V) \geq Z_{\beta\mu}^m(V_0) \qquad \text{(A.14)}$$

since by neglecting the corridor in the integral (A.5) a positive term in the partition sum is neglected. In conclusion we have demonstrated that

$$\frac{\log Z_{\beta\mu}(V)}{V} \underset{V \text{ and } V_0 \to \infty}{\longrightarrow} \frac{\log Z_{\beta\mu}(V_0)}{V_0}. \qquad \text{(A.15)}$$

It is very important to stress that this result is true only for short-range interactions. For these specific systems the implications of eq. (A.15) can be summarized as follows:

- A thermodynamic limit exists for these systems if the thermodynamic potential per unit volume tends to a constant for large volumes $\log Z_{\beta\mu}(V)/V \to \omega$;

– In the thermodynamic limit ensembles are equivalent. Indeed if $\omega = \log Z_{\beta\mu}(V_i)/V_i$ for an arbitrary subsystem V_i, using the fact that average values of extensive variables are first derivatives of $\log Z$ ($\langle A_\ell\rangle = -\partial_{\lambda_\ell}\log Z(\{\lambda_\ell\})$) and variances second derivatives ($\sigma_\ell^2 = \partial^2_{\lambda_\ell}\log Z(\{\lambda_\ell\})$, this implies that both are proportional to V_i. Then the average values per unit volume of extensive variables ($\rho_\ell = \langle A_\ell\rangle/V$) are independent of V and the variances of ρ_ℓ are inversely proportional to V, approaching zero as V goes to infinity. Since ensembles differ at the level of fluctuations, this demonstrates the equivalence between ensembles. For the explicit demonstration of the equality of the canonical and grand-canonical EoS we refer the reader to refs. [11,35].

References

1. T. Dauxois *et al.*, *Dynamics and Thermodynamics of Systems with Long Range Interactions*, Lect. Notes Phys., Vol. **602** (Springer, 2002).
2. D. Lynden-Bell, R. Wood, Mon. Not. R. Astron. Soc. **138**, 495 (1968); D. Lynden-Bell, Physica A **263**, 293 (1999).
3. R.M. Lynden-Bell, Mol. Phys. **86**, 1353 (1995).
4. J. Barré, D. Mukamel, S. Ruffo, Phys. Rev. Lett. **87**, 030601 (2001); J. Barré, F. Bouchet, T. Dauxois, S. Ruffo, J. Stat. Phys. **119**, 677 (2005); D. Mukamel, S. Ruffo, N. Schreiber, Phys. Rev. Lett. **95**, 240604 (2005).
5. T. Tatekawa, F. Bouchet, T. Dauxois, S. Ruffo, Phys. Rev. E **71**, 056111 (2005).
6. M. Schmidt *et al.*, Phys. Rev. Lett. **86**, 1191 (2001).
7. M. D'Agostino *et al.*, Phys. Lett. B **473**, 219 (2000).
8. F. Gobet *et al.*, Phys. Rev. Lett. **89**, 183403 (2002).
9. E. Melby *et al.*, Phys. Rev. Lett. **83**, 3150 (1999); S. Siem *et al.*, Phys. Rev. C **65**, 044318 (2002).
10. L.D. Landau, E.M. Lifshitz, *Statistical Physics* (Pergamon Press, 1980) Chapt. 3.
11. K. Huang, *Statistical Mechanics* (John Wiley and Sons Inc., 1963) Chapt. 5.
12. R.C. Tolman, *Principles of Statistical Mechanics* (Oxford University Press, London, 1962).
13. M. Rasetti, *Modern Methods in Statistical Mechanics* (World Scientific, Singapore, 1986).
14. E.T. Jaynes, *Information theory and statistical mechanics*, Stat. Phys., Brandeis Lect. **3**, 160 (1963).
15. D.H.E. Gross, *Microcanonical Thermodynamics: Phase transitions in Finite Systems*, Lect. Notes Phys., Vol. **66** (Springer, 2001).
16. R. Balian, *From Microphysics to Macrophysics* (Springer Verlag, 1982).
17. T.L. Hill, *Thermodynamics of Small Systems* (Dover, New York, 1994).
18. S. Abe, Y. Okamoto, *Nonextensive Statistical Mechanics and its Applications*, Lect. Notes Phys., Vol. **560** (Springer, 2001).
19. D.H.E. Gross, this topical issue and references therein.
20. F. Bouchet, J. Barré, J. Stat. Phys. **118**, 1073 (2005).
21. F. Leyvraz, S. Ruffo, Physica A **305**, 58 (2002).
22. F. Gulminelli, Ph. Chomaz, Phys. Rev. E **66**, 046108 (2002).
23. M. Costeniuc, R.S. Ellis, H. Touchette, J. Math. Phys. **46**, 063301 (2005).
24. M.S.S. Challa, J.H. Hetherington, Phys. Rev. Lett. **60**, 77 (1988); R.S. Johal, A. Planes, E. Vives, cond-mat/0307646.
25. W. Thirring, Z. Phys. **235**, 339 (1970).
26. A. Huller, Z. Phys. B **93**, 401 (1994).
27. R.S. Ellis, K. Haven, B. Turkington, J. Stat. Phys. **101**, 999 (2000).
28. T. Dauxois, P. Holdsworth, S. Ruffo, Eur. Phys. J. B **16**, 659 (2000).
29. J. Barré, D. Mukamel, S. Ruffo, Phys. Rev. Lett. **87**, 030601 (2001) cond-mat/0209357.
30. I. Ispolatov, E.G.D. Cohen, Physica A **295**, 475I (2001).
31. M. Kastner, M. Promberger, A. Huller, J. Stat. Phys. **99**, 1251 (2000); M. Kastner, J. Stat. Phys. **109**, 133 (2002); Physica A **359**, 447 (2006).
32. D.H.E. Gross, E.V. Votyakov, Eur. Phys. J. B **15**, 115 (2000).
33. A. Huller, M. Pleimling, Int. J. Mod. Phys. C **13**, 947 (2002).
34. M. Pleimling, H. Behringer, A. Huller, Phys. Lett. A **328**, 432 (2004); H. Behringer, J. Stat. Mech. P06014 (2005).
35. Ph. Chomaz, F. Gulminelli, in Lect. Notes Phys., Vol. **602** (Springer, 2002); F. Gulminelli, Ann. Phys. (Paris) **29**, 6 (2004).
36. P.H. Chavanis, I. Ispolatov, Phys. Rev. E **66**, 036109 (2002).
37. C.J. Pethick, H. Smith, *Bose Einstein Condensation in Dilute Gases* (Cambridge University Press, Cambridge, 2002).
38. A. Minguzzi *et al.*, Phys. Rep. **395**, 223 (2004).
39. E.V. Shuryak, Phys. Rep. **391**, 381 (2004).
40. P. Braun-Munzinger *et al.*, Phys. Lett. B **596**, 61 (2004); F. Becattini *et al.*, Phys. Rev. C **72**, 064904 (2005).
41. C. Brechignac *et al.*, Phys. Rev. Lett. **92**, 083401 (2004).
42. T.D. Lee, C.N. Yang, Phys. Rev. **87**, 404 (1952).
43. Ph. Chomaz, F. Gulminelli, V. Duflot, Phys. Rev. E **64**, 046114 (2001).
44. K.C. Lee, Phys. Rev. E **53**, 6558 (1996).
45. Ph. Chomaz, F. Gulminelli, Physica A **330**, 451 (2003).
46. K. Binder, D.P. Landau, Phys. Rev. B **30**, 1477 (1984).
47. P. Labastie, R.L. Whetten, Phys. Rev. Lett. **65**, 1567 (1990).
48. F. Gulminelli, Ph. Chomaz, V. Duflot, Europhys. Lett. **50**, 434 (2000).
49. H. Reinhardt, Nucl. Phys. A **413**, 475 (1984).
50. R. Balian, Y. Alhassid, H. Reinhardt, Phys. Rep. **131**, 1 (1986).
51. M. Costeniuc, R.S. Ellis, H. Touchette, B. Turkington, Phys. Rev. E **73**, 026105 (2006).
52. C. Menotti, P. Pedri, S. Stringari, Phys. Rev. Lett. **89**, 252402 (2002).
53. F. Gulminelli, Ph. Chomaz, Nucl. Phys. A **734**, 581 (2004); Ph. Chomaz, F. Gulminelli, O. Juillet, Ann. Phys. **320**, 135 (2005).
54. A.S. Botvina, I.N. Mishustin, this topical issue and references therein.
55. M. Di Toro, A. Olmi, R. Roy, this topical issue and references therein.
56. B. Tamain, this topical issue and references therein.
57. J.L. Mc Cauley, *Chaos Dynamics and Fractals*, Cambridge Nonlinear Science Series 2 (Cambridge University Press, 1993).

58. R. Balian, M. Vénéroni, Phys. Rev. Lett. **47**, 1353; 1765 (1981)(E); Ann. Phys. (N.Y.) **164**, 334 (1985).
59. Ph. Chomaz, Ann. Phys. (Paris) **21**, 669 (1996).
60. L. Van Hove, Physica **15**, 951 (1949); C.N. Yang, T.D. Lee, Phys. Rev. **87**, 404 (1952); K. Huang, *Statistical Mechanics* (John Wiley and Sons Inc., 1963) Chapt. 15.2 and appendix C.
61. P.H. Chavanis, M. Rieutord, Astron. Astrophys. **412**, 1 (2003); D.H.E. Gross, Entropy **6**, 158 (2004); A. De Martino, E.V. Votyakov, D.H.E. Gross, Nucl. Phys. B **654**, 427 (2003).
62. L. Casetti, M. Pettini, E.G.D. Cohen, Phys. Rep. **337**, 237 (2000).
63. I. Hidmi, D.H.E. Gross, H.R. Jaqaman, Eur. Phys. J. D **20**, 87 (2002).
64. T. Dauxois, V. Latora, A. Rapisarda, S. Ruffo, A. Torcini, in Lect. Notes Phys., Vol. **602** (Springer, 2002).
65. T.L. Beck, R.S. Berry, J. Chem. Phys. **88**, 3910 (1988); D.J. Wales, R.S. Berry, Phys. Rev. Lett. **73**, 2875 (1994); R.E. Kunz, R.S. Berry, J. Chem. Phys. **103**, 1904 (1995); D.J. Wales *et al.*, Adv. Chem. Phys. **115**, 1 (2001); R.S. Berry, Israel J. Chem. **44**, 211 (2004).
66. V.A. Antonov, Len. Univ. **7**, 135 (1962); IAU Symp. **113**, 525 (1995).
67. P. Hertel, W. Thirring, Ann. Phys. (N.Y.) **63**, 520 (1971).
68. P.H. Chavanis, in Lect. Notes Phys. Vol. **602** (Springer, 2002); Astron. Astrophys. **432**, 117 (2005).
69. T. Padhmanaban, in Lect. Notes Phys., Vol. **602** (Springer, 2002).
70. J. Katz, Not. R. Astron. Soc. **183**, 765 (1978).
71. M. Promberger, A. Huller, Z. Phys. B **97**, 341 (1995); **93**, 401 (1994).
72. H. Behringer, M. Pleimling, A. Hüller, J. Phys. A **38**, 973 (2005); H. Behringer, J. Phys. A **37**, 1443 (2004).
73. M. Kastner, M. Promberger, J. Stat. Phys. **53**, 795 (1988).
74. R. Franzosi, M. Pettini, L. Spinelli, Phys. Rev. E **60**, 5009 (1999); Phys. Rev. Lett. **84**, 2774 (2000).
75. R. Franzosi, M. Pettini, Phys. Rev. Lett. **92**, 60601 (2004); math-ph/0505057, math-ph/0505058.
76. J. Naudts, Europhys. Lett. **69**, 719 (2005) cond-mat/0412683.
77. M. Antoni, S. Ruffo, A. Torcini, Europhys. Lett. **66**, 645 (2004).
78. F. Gulminelli, Ph. Chomaz, A.H. Raduta, A.R. Raduta, Phys. Rev. Lett. **91**, 202701 (2003).
79. T.E. Strezelecka, M.W. Davidson, R.L. Rill, Nature **331**, 457 (1988); Y. Kafri, D. Mukamel, L. Peliti, Phys. Rev. Lett. **85**, 4988 (2000); A. Wynveen, D.J. Lee, A.A. Kornyshev, Eur. Phys. J. E **16**, 303 (2005).
80. F. Gulminelli, Ph. Chomaz, Phys. Rev. Lett. **82**, 1402 (1999).
81. Ph. Chomaz, F. Gulminelli, Nucl. Phys. A **647**, 153 (1999).
82. S. Grossmann, W. Rosenhauer, Z. Phys. **207**, 138 (1967); P. Borrmann *et al.*, Phys. Rev. Lett. **84**, 3511 (2000); H. Stamerjohanns *et al.*, Phys. Rev. Lett. **88**, 053401 (2002).
83. H. Touchette, Physica A **359**, 375 (2005).
84. M. Pichon, B. Tamain, R. Bougault, O. Lopez, Nucl. Phys. A **749**, 93 (2005).
85. O. Lopez *et al.*, Nucl. Phys. A **685**, 246 (2001).