Ensemble equivalence for general many-body systems

H. Touchette\(^{(a)}\)

School of Mathematical Sciences, Queen Mary University of London - London E1 4NS, UK

received 7 September 2011; accepted in final form 18 October 2011
published online 24 November 2011

PACS 05.20.Gg – Classical ensemble theory
PACS 05.70.-a – Thermodynamics
PACS 64.70.qd – Thermodynamics and statistical mechanics

Abstract – It has been proved for a class of mean-field and long-range systems that the concavity of the thermodynamic entropy determines whether the microcanonical and canonical ensembles are equivalent at the level of their equilibrium states, i.e., whether they give rise to the same equilibrium states. Here we show that this correspondence is actually a general result of statistical mechanics: it holds for any many-body system for which equilibrium states can be defined and in principle calculated. The same correspondence applies for other dual statistical ensembles, such as the canonical and grand-canonical ensembles.

Copyright © EPLA, 2011

Introduction. – The microcanonical ensemble (ME) and canonical ensemble (CE) provide two different statistical descriptions of many-body systems at equilibrium —the first in terms of their energy, the second in terms of the temperature of a heat bath connected to them. Although the two ensembles often give compatible or equivalent predictions in the thermodynamic limit, that equivalence is not guaranteed by their definitions. Gibbs noticed this point as early as 1902 [1] and, since then, many works have shown that the two ensembles either give equivalent or nonequivalent thermodynamic-limit predictions depending on the many-body system considered or, more precisely, on the type of interactions acting on the particles or constituents of that system (see [2] for a comprehensive survey).

The fundamental result that has emerged from these works is that many-body systems involving short-range and stable interactions, such as short-range spin systems or screened Coulomb systems, always have equivalent ensembles, whereas systems involving long-range interactions, such as gravitating systems, dipolar or unscreened Coulomb systems, may have nonequivalent ensembles [2]. How this equivalence or nonequivalence manifests itself depends on the level of description considered. At the thermodynamic level, the ME and CE become nonequivalent when the heat capacity is negative in the ME but is positive in the CE. This surprising phenomenon, which has been studied first in the context of gravitational systems [3–9], is now known to arise when the thermodynamic entropy of the ME is nonconcave as a function of the mean energy. More recently, it has been shown that the ME and CE can also be nonequivalent if they have different sets of equilibrium states —more precisely, if there exist equilibrium states in the ME, calculated at fixed energy, that do not arise in the CE at fixed temperature [10]. When this is the case, we say that the ME and CE are nonequivalent at the macrostate level. Various models have been shown to display this level of nonequivalence, including spin models [11–13] and models of 2D turbulence [14,15].

Given that thermodynamics is ultimately rooted in statistical mechanics, it is natural to seek for a link between the thermodynamic and macrostate levels of ensemble equivalence, i.e., to determine whether the concavity of the entropy determines the macrostate equivalence of ensembles. This problem was studied by Eyink and Spohn [16] and later by Ellis, Haven and Turkington [10], who proved that the ME and CE are, as physically expected, thermodynamically equivalent when they are equivalent at the macrostate level [17]. However, they were able to prove this result only for a limited class of systems that includes mostly mean-field–type systems.

The aim of this paper is to show that the connection between the thermodynamic and macrostate levels of ensemble equivalence is actually a general result of statistical mechanics: it holds for any many-body system for which equilibrium states can be defined and in principle calculated. This opens up the study of ensemble equivalence to many systems that do not fall in the class considered in [10] —most critically, gravitating systems, which are the source of much of the current research on long-range interactions [2,9]. This general result also implies that the ME

\(^{(a)}\)E-mail: ht@maths.qmul.ac.uk
and CE are always equivalent for any macrostates of any short-range systems having concave entropies and, more generally, any systems that do not have first-order phase transitions in the CE.

The rigorous proofs of the results stated in this paper will be published elsewhere [18]. Here we focus on presenting the physical meaning of these results, as well as the main ideas involved in proving them, following [10]. We illustrate the more general nature of our results by discussing the equivalence of the ME and CE for a simple spin model not completely covered by the results of [10]. We also derive, from our equivalence results, general results about phase coexistence and first-order phase transitions. Finally, we discuss their generalization to ensembles other than the ME and CE.

Framework. — We consider a general classical $N$-particle system described by the Hamiltonian or energy function $H(\omega)$, where $\omega \in \Lambda$ denotes the microstate of that system and $\Lambda$ the microstate space. The mean energy or energy per particle is denoted by $h(\omega) = H(\omega)/N$. We also consider a macrostate of that system, defined mathematically as a function $M(\omega)$ of the microstates. This macrostate can represent, for example, the magnetization of a spin system, the empirical distribution of velocities or positions of a gas of $N$ particles, or a combination of any observables defined similarly as functions of $\omega$.

The problem that we are concerned with is to define the set of equilibrium values of $M$ in both the ME and CE in the thermodynamic limit, and to determine whether these sets are equivalent, i.e., whether they can be put in some correspondence. Mathematically, the microcanonical equilibrium values of $M$, which we denote collectively as $E^u$, are identified as the typical values (viz., global maxima or concentration points) of the ME probability distribution $P^u(M = m)$ obtained by conditioning a uniform (prior) probability distribution $P$ on $\Lambda$ on the “constrained” set of microstates such that $h(\omega) = u$. In symbols, this is expressed as

$$P^u(m) = P(M = m|h = u) = \frac{P(M = m, h = u)}{P(h = u)}. \quad (1)$$

Note that, for simplicity, we use the same $P$ to denote the prior distribution $P(\omega)$ on $\Lambda$ and probabilities in general. Thus $P(M = m)$ denotes the probability (density) of the event $M = m$ calculated with respect to $P(\omega)$.

In a similar way, the canonical equilibrium values of $M$, denoted by $E^\beta$, are identified as the typical values of the CE probability distribution $P_\beta(M = m)$ of $M$ obtained by integrating the Gibbs distribution over the set of microstates such that $M(\omega) = m$. Thus,

$$P_\beta(M = m) = \int_{\omega \in \Lambda; M = m} P_\beta(\omega) \, d\omega, \quad (2)$$

where

$$P_\beta(\omega) = \frac{e^{-\beta H(\omega)}}{Z(\beta)}, \quad (3)$$

with

$$Z(\beta) = \int_{\Lambda} e^{-\beta H(\omega)} \, d\omega \quad (4)$$

as the partition function. The parameterization of the different distributions reflects of course the parameters defining each ensemble: the mean energy $u$ (or the energy $Nu$) for the ME, and the inverse temperature $\beta$ (or temperature $T = (k_B\beta)^{-1}$ with $k_B$ the Boltzmann constant) for the CE.

Mixture of ensembles. — It is obvious from the construction and physical interpretation of the ME and CE that $P^u(m)$ and $P_\beta(m)$ are in general two very different probability distributions of the same macrostate. The point about equivalence of ensembles, however, is that the typical values of $M$ determined by these distributions, which define, respectively, $E^u$ and $E_\beta$, may be equivalent in the sense that for a given $u$, there might be a given $\beta$ such that $E^u = E_\beta$. The essential idea involved in establishing this equivalence, which can be traced back to Gibbs [1], is that the CE is a probabilistic mixture of MEs.

To explain the meaning of this last statement, simply consider the Gibbs canonical distribution of eq. (3). Since this distribution depends only on $\beta$ and $H(\omega)$, it is clear that all microstates $\omega$ having the same energy have the same probabilistic weight, which implies that the conditional probability distribution $P_{\beta}(\omega|h = u)$, obtained by conditioning $P_\beta(\omega)$ on the set of microstates such that $h(\omega) = u$, must be uniform over that constrained set of microstates. Since $P^u(\omega)$ has exactly this property, we must therefore have $P_\beta(\omega|u) = P^u(\omega)$ for all $\omega \in \Lambda$ and by extension $P_\beta(m|u) = P^u(m)$.

With this result, we can use Bayes’s Theorem to write

$$P_\beta(m) = \int_{\mathbb{R}} P_\beta(m|u) \, P_\beta(u) \, du = \int_{\mathbb{R}} P^u(m) \, P_\beta(u) \, du, \quad (5)$$

where $P_\beta(u)$ is the probability distribution of the mean energy $h$ in the CE. In terms of the density of states $\Omega(u)$, this distribution is simply given by

$$P_\beta(u) = \frac{\Omega(u) \, e^{-\beta Nu}}{Z(\beta)}, \quad (6)$$

The second equality obtained in (5) shows that the CE is a superposition of MEs weighted by the canonical probability distribution $P_\beta(u)$ of the mean energy. It is this superposition, which is valid for all $N$, that we refer to as a probabilistic mixture of MEs.

Results. — The idea that the CE is a superposition of MEs is simple but far reaching because it implies, by taking the thermodynamic limit, that the concentration points of $P_\beta(m)$ are the concentration points of $P^u(m)$ where $P_\beta(u)$ concentrates. In other words, $E_\beta$ must be given by all the sets $E^u$ such that $u$ is an equilibrium value of the mean energy in the CE at inverse temperature $\beta$. To put this statement in mathematical form, let us denote by $U_\beta$ the set of these mean energies realized at equilibrium in the CE at $\beta$. Then we obtain from
The physical interpretation of this case is that the equilibrium states obtained in the ME for values of $u$ vertex model of 2D turbulence [14], field Blume-Emery-Griffiths model [12], and the point-vertex model of 2D turbulence [14].

**Case 3 (Partial equivalence):** If $s$ is concave but not strictly concave at $u$ (point $c$ in fig. 1), then $u \in \mathcal{U}_\beta$ for some $\beta \in \mathbb{R}$, but $u$ is not the only member of $\mathcal{U}_\beta$, which implies from eq. (5) that $\mathcal{E}^u \subseteq \mathcal{E}_\beta$.

This last result follows because $\mathcal{E}_\beta$ in this case is made up of different microcanonical equilibrium states at different mean energies, one of which is $u$. Thus $\mathcal{E}_\beta$ contains all the equilibrium states of the ME at $u$ but may also contain more equilibrium states not seen in $\mathcal{E}^u$; hence the name “partial equivalence”. If any of the many sets that compose $\mathcal{E}_\beta$ is different from $\mathcal{E}^u$, which is expected since different energies should give rise to different equilibrium states, then the previous result is strengthened to $\mathcal{E}^u \subset \mathcal{E}_\beta$. That is, in this case $\mathcal{E}_\beta$ does contain equilibrium states that are not contained in $\mathcal{E}^u$.

**Relationship with phase transitions.** – Partial equivalence of ensembles is a boundary case between the equivalent and nonequivalent cases, which corresponds physically to a phase coexistence accompanying a first-order phase transition in the CE [17]. To understand this point, consider the two points labelled by $c$ in fig. 1. These two points are not strictly concave because they share the same supporting line, which corresponds to the line of the concave envelope $s^{**}(u)$ of $s(u)$. Let us denote the mean energies of these points by $u_l$ and $u_h$, respectively, and let $\beta_c$ be the slope of their supporting line. The non-strict concavity of these points then implies that $F_{\beta_c}(u)$ is a bimodal distribution with two equal-height peaks at $u_l$ and $u_h$ [17], so that $\mathcal{U}_{\beta_c} = \{u_l, u_h\}$ and $\mathcal{E}_{\beta_c} = \mathcal{E}^{u_l} \cup \mathcal{E}^{u_h}$. This shows that the CE at $\beta_c$ is composed of many equilibrium states of the ME. If we regard each of these microcanonical sets as a “pure phase” of the ME, then we can say that the CE shows phase coexistence at $\beta_c$.

This result is consistent with the fact that the CE shows a first-order phase transition at $\beta_c$. Indeed, it is known that if the entropy $s(u)$ is nonconcave, for example as in fig. 1, then the canonical free energy,

$$\varphi(\beta) = \lim_{N \to \infty} \frac{1}{N} \ln Z(\beta),$$

is nondifferentiable at some point, corresponding in fig. 1 to $\beta_c$. We thus see that the set of equilibrium states of any system in the CE is, at a point of first-order phase transition in the CE, the combination of two or more sets or “phases” of equilibrium states of the ME. This does not imply that each equilibrium state in the CE is itself a superposition or a mixture of two or more equilibrium states of the ME; only that the whole set of equilibrium states in the CE is a such a mixture.

Note, finally, that the connection between nonconcave entropies and nondifferentiable free energies also goes in the opposite direction, in that, if $\varphi(\beta)$ is everywhere differentiable, then $s(u)$ is necessarily strictly concave. We therefore obtain as a general physical result that the ME and CE are equivalent at the macrostate level for any systems that do not have first-order phase transitions.
Comparison with previous results. – The result of eq. (5) and the three cases stated above were first obtained, as mentioned before, by Ellis, Haven and Turkington [10]. The contribution of the present paper is to re-derive these results under much more general conditions than originally done. All that is required here is for the distributions \( P^\alpha(m) \) and \( P^\beta(m) \) to exist, so that the equilibrium sets \( E^\alpha \) and \( E^\beta \) can be constructed. This is a very weak assumption, which amounts essentially to saying that the results of [10] actually hold provided that the thermodynamic limit of the ME and CE is well defined. In fact, it can be proved that the existence of \( E^\alpha \) and \( E^\beta \) is essentially equivalent to the existence of \( s(u) \) and \( \varphi(\beta) \), respectively, so that our results about the equivalence and nonequivalence of the ME and CE simply hold provided these thermodynamic functions exist.

In [10], the same results were obtained by comparison under two technical conditions. The first is that the mean energy function \( h(\omega) \) can be rewritten as a function of some macrostate \( M(\omega) \) in the thermodynamic limit. Mathematically, this amounts to saying that there exists a function \( \tilde{h}(m) \) of the macrostate \( M \), called the energy representation function, such that

\[
\lim_{N \to \infty} |h(\omega) - \tilde{h}(M(\omega))| = 0
\]

uniformly for all \( \omega \in \Lambda \). The second condition is that there exists an entropy function \( \tilde{s}(m) \) associated with \( M \). This macrostate entropy is defined similarly as \( s(u) \) by replacing the density of states \( \Omega(u) \) with the density of states \( \Omega(m) \) associated with the microstates having a given macrostate value \( M(\omega) = m \) (see [10] for the precise definition).

These two conditions limit greatly the type of systems for which the equivalence of the ME and CE could be analyzed. The first condition in particular is akin to a mean-field requirement, known to be satisfied for mean-field systems and certain long-range interacting systems [10,12,19]. It does not apply in a uniform sense to gravitational systems, among other long-range systems, and certainly not for systems with short-range interactions. Moreover, even for mean-field or long-range systems, there is typically only one macrostate that qualifies as a macrostate for which \( \tilde{h} \) and \( \tilde{s} \) exist, which means that the equivalence of the ME and CE can only be discussed for this particular macrostate.

In our version of the equivalence and nonequivalence results, these restrictions are completely lifted: we can discuss the equivalence or nonequivalence of the ME and CE for any systems and any macrostates or observables provided that the equilibrium values of these observables can be defined in each ensemble in the thermodynamic limit. When this is the case, the ME and CE are equivalent essentially if and only if the microcanonical entropy is concave in the thermodynamic limit as a function of the mean energy.

Application. – To illustrate the generality of our results, compared to those of [10], we now discuss the equivalence of the ME and CE for the \( \alpha \)-Ising model, defined by the Hamiltonian

\[
H = \frac{J}{N^{1-\alpha}} \sum_{i \neq j=1}^{N} \frac{1 - S_i S_j}{|i - j|^\alpha},
\]

where \( S_i = \pm 1 \) is a spin variable. This model has been extensively studied because of its rich phase transition behavior [20–23], which includes a suspected Kosterlitz-Thouless transition for \( \alpha = 2 \) [24,25]. For \( J > 0 \) and \( 0 \leq \alpha < 1 \), Barré et al. [19] were able to show that the model admits an energy representation function \( \tilde{h} \) and macrostate entropy \( \tilde{s} \) for a particular macrostate, namely, the macroscopic magnetization profile \( m(x) \) (see [19] for the precise definition of this macrostate). Given that the entropy \( s(u) \) is concave as a function of \( u \) for these parameters, they then concluded that the ME and CE are equivalent at the level of \( m(x) \).

For \( \alpha \geq 1 \), there is no known energy representation function, so the equivalence of the ME and CE cannot be discussed for this case using the results of [10]. However, the entropy \( s(u) \) exists for these parameters and is everywhere concave, since it is known that the model does not show any first-order phase transitions. From our results, we therefore conclude that the ME and CE are equivalent for this model at the level of \( m(x) \). More importantly, they are also equivalent at the level of any macrostate, such as the standard magnetization or the spin distribution, which do not satisfy the requirements of [10]. Similar conclusions can be reached for other models and macrostates, including short-range models, such as the Ising model at the level of its magnetization, and long-range models such as self-gravitating particle models [18].

Idea of the proof. – We conclude by giving an overview of how the results of this paper are proved at a rigorous level. The first step, following [10], is to use large deviation theory [26] to derive a large deviation principle (LDP) for each of the distributions \( P^\alpha(m) \) and \( P^\beta(m) \), and to use these LDPs to define (in the thermodynamic limit) the sets \( \mathcal{E}^\alpha \) and \( \mathcal{E}^\beta \) of equilibrium values of the macrostate \( M \) in the ME and CE, respectively.

Heuristically, we express these LDPs by the approximations

\[
P^\alpha(m) \approx e^{-N I^\alpha(m)}
\]

and

\[
P^\beta(m) \approx e^{-N I^\beta(m)}
\]

which are assumed to hold with sub-exponential correction terms in the limit \( N \to \infty \). Given the rate functions \( I^\alpha(m) \) and \( I^\beta(m) \), the sets \( \mathcal{E}^\alpha \) and \( \mathcal{E}^\beta \) are defined, respectively, as the sets of global minima of \( I^\alpha(m) \) and \( I^\beta(m) \) [10].

Next, we use the idea that the CE is a probabilistic mixture of MEs, i.e., eq. (5), to obtain the following large
deviation approximation:

\[ P_{\beta}(m) \approx e^{-N(J^\beta(u))} \]

which involves the rate function \( J^\beta(u) \) of \( P^\beta(m) \) and the rate function \( J_\beta(u) \) associated with the LDP of the distribution \( P_\beta(u) \) of the mean energy \( h \) in the CE:

\[ P_\beta(u) \approx e^{-N J_\beta(u)}. \]

In terms of \( s(u) \) and \( \varphi(\beta) \), \( J_\beta(u) \) is explicitly given by [26]:

\[ J_\beta(u) = \beta u - s(u) - \varphi(\beta). \]

By heuristically approximating the integral (14) using Laplace’s method, we are then led to the following representation formula connecting the canonical and microcanonical rate functions:

\[ I_\beta(m) = \inf_{u \in \mathcal{R}} \{ I^\beta(m) + J_\beta(u) \}. \]  

Interestingly, this representation formula connects not only the equilibrium states of the ME and CE, but also the fluctuations of \( M \) around these equilibrium states in each ensemble via the fluctuations of \( h \) in the CE.

With the formula (17) and the explicit expression of \( J_\beta(u) \), we finally arrive at eq. (5) by noting that \( U_\beta \) is the set of global minima of \( J_\beta(u) \) and, from there, at the three cases of equivalence mentioned before by studying these global minima in relation to the concavity of \( s(u) \). We leave the details of this last step, which involves results from convex analysis, and the rigorous justification of all the approximations mentioned above to a more technical paper [18].

Summary. – We have shown in this paper that the equivalence of the ME and CE at the level of their equilibrium states is determined by the concavity of the microcanonical entropy \( s(u) \) in a model-independent way. Such a relationship between the thermodynamic and macrostate properties of equilibrium systems was known to hold for mean-field and some long-range systems. Here we have shown that it holds for any many-body system and any observable, so long as their equilibrium behavior is well defined. The same relationship can be derived for other ensembles that are dual in the sense of the ME and CE (e.g., the CE and grand-canonical ensemble or the fixed-magnetization and fixed-magnetic field ensembles) simply by replacing \( E^\beta, E_\beta \) and \( s(u) \) with the appropriate sets of equilibrium states and entropy, respectively. The entropy determining the equivalence of the CE and grand-canonical ensemble, for example, is the entropy \( s(\rho) \) expressed as a function of the particle density \( \rho \). Finally, our results can also be generalized to nonequilibrium ensembles (see, e.g., [26–28]), as well as quantum systems when observables that commute with the energy are considered [29].

I wish to thank R. Nerattini, C. Nardini, G. L. Sewell and R. J. Harris for useful comments. This work was partially supported by an Interdisciplinary Fellowship from RCUK.

REFERENCES

[1] Gibb J. W., Elementary Principles in Statistical Mechanics with Special Reference to the Rational Foundation of Thermodynamics (Yale University Press) 1902.
[2] Campa A., Dauxois T. and Ruffo S., Phys. Rep., 480 (2009) 57.
[3] Lynden-Bell D. and Wood R., Mon. Not. R. Astron. Soc., 138 (1968) 495.
[4] Lynden-Bell D. and Lynden-Bell R. M., Mon. Not. R. Astron. Soc., 181 (1977) 405.
[5] Lynden-Bell D., Physica A, 263 (1999) 293.
[6] Hertel P. and Thyring W., Ann. Phys. (N. Y.), 63 (1971) 520.
[7] Gross D. H. E., Microcanonical Thermodynamics: Phase Transitions in “Small” Systems, Lect. Notes Phys., Vol. 66 (World Scientific, Singapore) 2001.
[8] Ispolatov I. and Cohen E. G. D., Phys. Rev. Lett., 87 (2001) 210601.
[9] Chavanis P.-H., Int. J. Mod. Phys. B, 20 (2006) 3113.
[10] Ellis R. S., Haven K. and Turington B., J. Stat. Phys., 101 (2000) 999.
[11] Barré J., Mekam D. and Ruffo S., Phys. Rev. Lett., 87 (2001) 030601.
[12] Ellis R. S., Touchette H. and Turington B., Physica A, 335 (2004) 518.
[13] Costeniucl M., Ellis R. S. and Touchette H., J. Math. Phys., 46 (2005) 063301.
[14] Ellis R. S., Haven K. and Turington B., Nonlinearity, 15 (2002) 239.
[15] Venaille A. and Boucl C., Phys. Rev. Lett., 102 (2009) 104501.
[16] Eyink G. L. and Spohn H., J. Stat. Phys., 70 (1993) 835.
[17] Touchette H., Ellis R. S. and Turington B., Physica A, 340 (2004) 138.
[18] Touchette H., in preparation (2011).
[19] Barré J., Boucl C., Dauxois T. and Ruffo S., J. Stat. Phys., 119 (2005) 677.
[20] Dyson F. J., Commun. Math. Phys., 12 (1969) 91.
[21] Fisher M. E., Ma S.-K. and Nickel B. G., Phys. Rev. Lett., 29 (1972) 917.
[22] Fröhlich J. and Spencer T., Commun. Math. Phys., 84 (1982) 87.
[23] Luften E. and Bléte H. W. J., Phys. Rev. B, 56 (1997) 8945.
[24] Trouless D. J., Phys. Rev., 187 (1969) 732.
[25] Kosterlitz J. M., Phys. Rev. Lett., 37 (1976) 1577.
[26] Touchette H., Phys. Rep., 478 (2009) 1.
[27] Evans R. M. L., J. Phys. A: Math. Gen., 38 (2005) 293.
[28] Jack R. L. and Sollich P., Prog. Theor. Phys. Suppl., 184 (2010) 304.
[29] Kastner M., Phys. Rev. Lett., 104 (2010) 240403.