Physica A 505 (2018) 864–870

Contents lists available at ScienceDirect

Physica A

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

Temperature is not an observable in superstatistics

Sergio Davis [a,b,*], Gonzalo Gutiérrez [c]

[a] Comisión Chilena de Energía Nuclear, Casilla 188-D, Santiago, Chile
[b] Departamento de Física, Facultad de Ciencias Exactas, Universidad Andres Bello, Sazié 2212, piso 7, 8370136, Santiago, Chile
[c] Grupo de Nanomateriales, Departamento de Física, Facultad de Ciencias, Universidad de Chile, Casilla 653, Santiago, Chile

HIGHLIGHTS

- Superstatistical temperature distributions cannot be recovered in general by any microscopically observable estimator.
- Energy and temperature are not in the same footing as thermodynamical quantities when superstatistics is considered.
- Superstatistics is best understood as Bayesian thermodynamics with a unique but uncertain temperature.

ARTICLE INFO

Article history:
Received 10 October 2017
Received in revised form 5 February 2018
Available online 11 April 2018

Keywords:
superstatistics

ABSTRACT

Superstatistics (Beck and Cohen, 2003) is a formalism that attempts to explain the presence of distributions other than the Boltzmann–Gibbs distributions in Nature, typically power-law behavior, for systems out of equilibrium such as fluids under turbulence, plasmas and gravitational systems. Superstatistics postulates that those systems are found in a superposition of canonical ensembles at different temperatures, and sometimes the physical interpretation is one of local thermal equilibrium in the sense of an inhomogeneous temperature distribution in different regions of space or instants of time.

Here we show that, in order for superstatistics to be internally consistent, it is impossible to define a phase-space function or microscopic observable $B(\boldsymbol{p}, \boldsymbol{q})$ corresponding one-to-one to the local value of $\beta = 1/k_B T$. Thus, unlike energy which is defined by a phase-space function $\mathcal{H}(\boldsymbol{p}, \boldsymbol{q})$ (the Hamiltonian), temperature is not a microscopic observable.

An important consequence of our proof is that, in Superstatistics, the identification of temperature with the kinetic energy is limited to the expectation of β and cannot be used to measure the different temperatures in local thermal equilibrium or its fluctuations.

1. Introduction

Superstatistics [1,2] is a relatively new, but already widely used [3–6] formalism which attempts to explain the appearance of non-Boltzmann distributions in Nature for non-equilibrium steady-state systems [7]. It postulates a weighted superposition of canonical ensembles at different temperatures, thus allowing the existence of temperature fluctuations around its average. In fact, in part due to the appeal of superstatistics, the long-held discussion about temperature fluctuations in thermodynamics [8,9] has led to a resurgence of interest particularly for the statistical mechanics of small systems [10–13] and also because fluctuations of β may be connected to the non-extensivity parameter q in Tsallis statistics [1].

* Corresponding author at: Comisión Chilena de Energía Nuclear, Casilla 188-D, Santiago, Chile.
E-mail addresses: sdavis@cchen.cl (S. Davis), gonzalogutierrez@uchile.cl (G. Gutiérrez).
URLs: http://www.lpmd.cl/sdavis (S. Davis), http://www.gnm.cl/gonzalo (G. Gutiérrez).

https://doi.org/10.1016/j.physa.2018.04.007
0378-4371/

Superstatistics has the advantage of not requiring a generalization of the entropy functional such as Tsallis' entropy [14]; it is based solely on the canonical ensemble and the correct application of the laws of probability [15]. In principle, superstatistics only deals with uncertain values of a parameter β. However there is a somewhat implicit assumption that fluctuating temperatures are measurable (for instance, in atomistic computer simulation) if temperature is essentially the kinetic energy per particle. It seems natural to associate the fluctuations of kinetic energy, $\left\langle(\delta K)^2\right\rangle$, to the fluctuations of temperature, $\left\langle(\delta\hat{T}_K)^2\right\rangle$, where

$$T_K(\boldsymbol{p}) = \frac{2K}{3Nk_B} \tag{1}$$

is the kinetic estimator of temperature for a system with Hamiltonian $\mathcal{H}(\boldsymbol{r},\boldsymbol{p}) = K(\boldsymbol{p}) + \Phi(\boldsymbol{r})$. But, as is discussed by Falcioni et al. in Ref. [11], this identification is in fact incorrect for the fluctuations. In general, there are many possible definitions of *estimators of temperature*. As shown by Rugh [16] and Rickayzen [17], the inverse temperature $\beta = 1/k_BT$ can be obtained as the average of any quantity of the form

$$\hat{\beta}(\boldsymbol{r},\boldsymbol{p}) = \nabla\cdot\left[\frac{\boldsymbol{\omega}}{\boldsymbol{\omega}\cdot\nabla\mathcal{H}}\right], \tag{2}$$

where $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{r},\boldsymbol{p})$ is a differentiable vector function of positions and/or momenta, and the ∇ operator acts on both $\boldsymbol{r}$ and $\boldsymbol{p}$. This definition of $\hat{\beta}(\boldsymbol{r},\boldsymbol{p})$, known as dynamical temperature, was originally derived in the microcanonical ensemble but is completely general, independent of the statistical ensemble [18]. It allows us to write the microscopically observable temperature in terms of any combination of positions and/or momenta. For instance, a purely configurational inverse temperature function can be obtained as

$$\hat{\beta}_C(\boldsymbol{r}) = \nabla\cdot\left[\frac{\boldsymbol{\omega}}{\boldsymbol{\omega}\cdot\nabla\Phi}\right] \tag{3}$$

where $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{r})$ is a function of position.

The question remains: if $\left\langle(\delta\hat{T}_K)^2\right\rangle \neq \left\langle(\delta T)^2\right\rangle$, is this a shortcoming of the kinetic estimator $\hat{T}_K$? Is there another estimator $\hat{T}$ whose fluctuations agree with the "true" fluctuations of temperature? In current and future applications of Superstatistics to computer simulation of steady-state systems this question gains even more importance, as the inverse temperature has a full-fledged probability distribution $P(\beta|S)$. Is it possible to interpret $P(\beta|S)$ as a frequency distribution of a microscopically observable inverse temperature $\hat{\beta}$, a function of the coordinates and momenta of the system? See for instance Fig.1 in Ref. [19] in which a schematic of a superstatistical system is depicted, where different regions of space have different inverse temperatures β_1, β_2 and so on. Through such a function $\hat{\beta}$ one could in principle reveal the distribution of inverse temperatures $P(\beta|\mathcal{S})$ of a system by making a histogram.

In this work, we show that in general for superstatistics (outside the trivial case with zero fluctuations of β, where superstatistics reduces to a thermodynamic limit ensemble such as the canonical) there is no microscopically observable function $B(\boldsymbol{r},\boldsymbol{p})$ such that the frequency distribution of B matches $P(\beta|S)$. That is, not only the kinetic estimator temperature T_K of Eq. (1) fails in its role as a measure of local or instantaneous temperature in superstatistical systems but the problem is deeper: the goal of finding a function $T(\boldsymbol{r},\boldsymbol{p})$ with a one-to-one correspondence to the value of the "true" temperature T (in the sense of the probability density of β) cannot be achieved.

The paper is organized as follows. In Section 2, a few elements of Statistical Mechanics are reviewed, mainly to establish the notation. Then in Section 3 these ideas are extended to ensembles with arbitrary fluctuations of energy. Section 4 presents the problem of inferring the underlying ensemble from a set of measurements, and it is in this context that a hypothetical phase-space function associated to temperature is postulated. Section 5 follows with the proof of impossibility of that function. Finally, Section 6 closes with some conclusions.

2. The framework of statistical mechanics

Consider a system with degrees of freedom $\boldsymbol{\Gamma} = (\boldsymbol{r},\boldsymbol{p})$ and Hamiltonian $\mathcal{H}(\boldsymbol{\Gamma})$, whose values we will denote by E. This Hamiltonian is bounded from below but not from above, i.e., $E_0 < \mathcal{H}(\boldsymbol{\Gamma}) < \infty$. The minimum energy E_0 can be set to zero without loss of generality. If the system is perfectly isolated so that its energy is strictly fixed at a value E, the probability distribution of the different microstates is given by the microcanonical ensemble [20],

$$P(\boldsymbol{\Gamma}|E,V,N) = \frac{1}{\Omega(E;V,N)}\delta(\mathcal{H}(\boldsymbol{\Gamma})-E), \tag{4}$$

where

$$\Omega(E;V,N) = \int d\boldsymbol{\Gamma}\,\delta(\mathcal{H}(\boldsymbol{\Gamma})-E) \tag{5}$$

is the density of states. If, on the other hand, the system is placed inside a heat bath at temperature T, the probability distribution of the states is the canonical ensemble,

$$P(\boldsymbol{\Gamma}|\beta) = \frac{\exp\left(-\beta\mathcal{H}(\boldsymbol{\Gamma})\right)}{Z(\beta)}. \tag{6}$$

with $\beta = 1/k_B T$ and

$$Z(\beta) = \int d\mathbf{\Gamma} \exp\left(-\beta\mathcal{H}(\mathbf{\Gamma})\right) = \int_0^\infty dE\,\Omega(E)\exp(-\beta E) \tag{7}$$

the partition function. In order for $Z(\beta)$ to be well-defined, the temperature T (and therefore β) cannot be negative. This temperature, in turn, can be connected with the density of states through the relation

$$\frac{1}{T} = \frac{\partial S(E; V, N)}{\partial E} \tag{8}$$

with $S(E; V, N) = k_B \ln \Omega(E; V, N)$ the Boltzmann entropy.

We also know temperature is related to the average kinetic energy of the system through the equipartition theorem,

$$\left\langle \sum_{i=1}^{N} \frac{p_i^2}{2m_i} \right\rangle_\beta = \frac{3N}{2} k_B T, \tag{9}$$

where $\langle\cdot\rangle_\beta$ denotes an expectation taken over the canonical distribution with given β.

3. Non-canonical stationary states

Let us now assume we place the system in a macroscopic stationary state $\mathcal{S}$, which is neither perfectly isolated nor in equilibrium with a heat bath. In this case, energy will fluctuate with $\langle(\delta E)^2\rangle_\mathcal{S} > 0$ and following a probability distribution $P(E|\mathcal{S})$. We can always describe the new distribution of microstates $P(\mathbf{\Gamma}|\mathcal{S})$ as a superposition of microcanonical ensembles weighted by $P(E|\mathcal{S})$, that is,

$$P(\mathbf{\Gamma}|\mathcal{S}) = \int_0^\infty dE P(E|\mathcal{S}) P(\mathbf{\Gamma}|E). \tag{10}$$

Replacing the definition of the microcanonical ensemble (Eq. (4)), we obtain

$$P(\mathbf{\Gamma}|\mathcal{S}) = \int_0^\infty dE \left[\frac{P(E|\mathcal{S})}{\Omega(E)}\right] \delta(\mathcal{H}(\mathbf{\Gamma}) - E) = \rho(\mathcal{H}(\mathbf{\Gamma})), \tag{11}$$

where we have defined, for simplicity of notation, the function $\rho(E)$ such that $P(E|\mathcal{S}) = \rho(E)\Omega(E)$. We see that the probability distribution of the microstates is a function of the Hamiltonian only, as required by the stationary Liouville equation,

$$\left\{P(\mathbf{\Gamma}|\mathcal{S}), \mathcal{H}(\mathbf{\Gamma})\right\} = \left\{\rho(\mathcal{H}(\mathbf{\Gamma})), \mathcal{H}(\mathbf{\Gamma})\right\} = 0. \tag{12}$$

In this case, unlike the microcanonical and canonical ensembles, the ensemble cannot be described by a single number such as E or β, instead it can only be described completely if we know the shape of the **function** ρ; In this sense we can say that it is, in fact, a statistical model with an infinite number of parameters.

An alternative to the decomposition in Eq. (10) is superstatistics, where $P(\mathbf{\Gamma}|\mathcal{S})$ is expressed as a superposition of canonical ensembles with different values of β, that is,

$$P(\mathbf{\Gamma}|\mathcal{S}) = \int_0^\infty d\beta P(\beta|\mathcal{S}) P(\mathbf{\Gamma}|\beta). \tag{13}$$

Replacing the definition of the canonical ensemble (Eq. (6)) and calling $E = \mathcal{H}(\mathbf{\Gamma})$ we have

$$\rho(E) = \int_0^\infty d\beta \left[\frac{P(\beta|\mathcal{S})}{Z(\beta)}\right] \exp(-\beta E) \tag{14}$$

from which we see that $\rho(E)$ is the Laplace transform of a new function $f(\beta)$ such that $P(\beta|\mathcal{S}) = f(\beta)Z(\beta)$. This means the function $f(\beta)$ also contains a full description of the macrostate $\mathcal{S}$, and for this purpose a determination of $f(\beta)$ is equivalent to a determination of $\rho(E)$. We will call these functions the *ensemble functions*.

It is important to emphasize here the fact that $f(\beta)$ does not correspond to the probability of observing values of β, in the same way that $\rho(E)$ is not the probability of observing the energy E. This has somewhat led to confusion in the literature. The connection between these *ensemble functions* f, ρ and the probability distributions $P(\beta|\mathcal{S})$ and $P(E|\mathcal{S})$ is given by the partition function and density of states, respectively. A brief summary of this information is given in Table 1.

4. Can we deduce the stationary ensemble from phase-space measurements?

Suppose that we have access to measurements of energy for a particular system in a stationary state, and we wish to determine the function ρ. We proceed to sample n values of energy $E_1, E_2, \ldots, E_n$ and construct an histogram h, as

$$h_j = \frac{1}{n}\sum_{i=1}^{n} \delta(j, k(E_i)) \tag{15}$$

Table 1
Features of a superstatistical stationary state $\mathcal{S}$. Note that our main result finally shows that there is no suitable definition of the function $B(\mathbf{\Gamma})$.

Property	Observable	Ensemble function	Probability density
E	$\mathcal{H}(\mathbf{\Gamma})$	$\rho(E)$	$\rho(E)\Omega(E)$
β	$B(\mathbf{\Gamma})$	$f(\beta)$	$f(\beta)Z(\beta)$

where $\delta(j, k)$ is Kronecker's delta, $k(E)$ gives the integer position of the bin corresponding to the value of energy E, and $j = 1, 2, \ldots, m$ with m the total number of bins.

If n and m are sufficiently large, by the law of large numbers

$$h_j \rightarrow \Big\langle \delta(E_j - \mathcal{H}(\mathbf{\Gamma}_i)) \Big\rangle_{\mathcal{S}} = P(E_j|\mathcal{S}) \tag{16}$$

i.e., the histogram will converge to the energy probability distribution $P(E|\mathcal{S})$, and so, in practice, we can obtain $\rho(E)$ from a large number of energy measurements if we know the density of states, as

$$\frac{h_j}{\Omega(E_j)} \approx \rho(E_j). \tag{17}$$

If we numerically obtain $\rho(E)$ in this way, we could apply the inverse Laplace transform and recover the ensemble function $f(\beta)$. But this is redundant because in that case we already would have $\rho(E)$, which has all the information to describe the system. We would like a more direct route to obtain $f(\beta)$, and then the following question arises:

Is β the value of a phase-space function $B(\mathbf{\Gamma})$ in the same way that E is the value of the Hamiltonian $\mathcal{H}(\mathbf{\Gamma})$?

If such a quantity B exists, and we know the partition function, we can directly obtain $f(\beta)$ without the intermediate step of computing $\rho(E)$, just by accumulating enough samples $\beta_1 = B(\mathbf{\Gamma}_1)$, $\beta_2 = B(\mathbf{\Gamma}_2)$, $\ldots$, $\beta_n = B(\mathbf{\Gamma}_n)$ and the relation

$$\frac{b_j}{Z(\beta_j)} \approx f(\beta_j), \tag{18}$$

analogous to Eq. (17), where now b_j is the histogram of values β_i, for which the law of large numbers holds as

$$b_j \rightarrow \Big\langle \delta(B(\mathbf{\Gamma}_i) - \beta_j) \Big\rangle_{\mathcal{S}} \tag{19}$$

and that we can identify with the probability distribution of β by

$$P(\beta|\mathcal{S}) = \Big\langle \delta(B(\mathbf{\Gamma}) - \beta) \Big\rangle_{\mathcal{S}}. \tag{20}$$

This is equivalent to the strong requirement that, for any test function $g(\beta)$,

$$\Big\langle g(\beta) \Big\rangle_{S} = \Big\langle g(B(\mathbf{\Gamma})) \Big\rangle_{\mathbf{S}}. \tag{21}$$

5. Impossibility of an intrinsic phase-space function for $\boldsymbol{\beta}$

In classical statistical mechanics, we expect that the microscopic observables O in our system are defined as phase-space functions $O(\mathbf{\Gamma})$ which are independent of the external conditions, being at most functionals of the Hamiltonian (which contains all the information about the system and its dynamics). In particular, we expect that if we place the system in a stationary *ensemble* $\mathcal{S}$, the **definition** of the observable, $O(\mathbf{\Gamma})$, will not change, despite the fact that its **value** $\big\langle O \big\rangle_{\mathcal{S}}$ most probably will. That is, we expect that O is not dependent on the *ensemble* function ρ. This condition can be expressed as

$$\frac{\delta O(\mathbf{\Gamma})}{\delta \rho(E)} = 0. \tag{22}$$

We will call the observables for which this is true, *intrinsic* observables. They can be defined "once and for all" if we know the Hamiltonian of the system.

Our main result is that β does not fall into this category: there is no intrinsic observable $B(\mathbf{\Gamma})$ which gives the superstatistical β, as shown by the following theorem.

Theorem. *In superstatistics, there is no phase-space function $B(\mathbf{\Gamma})$ such that*

$$P(\beta|\mathcal{S}) = \Big\langle \delta(B(\mathbf{\Gamma}) - \beta) \Big\rangle_{\mathcal{S}},$$

and

$$\frac{\delta B}{\delta \rho(E)} = 0.$$

That is, $B(\boldsymbol{\Gamma})$ is not an intrinsic observable of the system: even worse, its definition is dependent on the external conditions that maintain the stationary state, and thus cannot be used to infer the *ensemble*. In other words, every stationary ensemble $\mathcal{S}$ would have its own microscopic definition of temperature.

Proof. Replacing Eq. (20) into Eq. (13), we see that

$$\rho(\mathcal{H}(\boldsymbol{\Gamma})) = \int d\boldsymbol{\Gamma}' \rho(\mathcal{H}(\boldsymbol{\Gamma}'))\frac{\exp\left(-B(\boldsymbol{\Gamma}')\mathcal{H}(\boldsymbol{\Gamma})\right)}{Z(B(\boldsymbol{\Gamma}'))}. \tag{23}$$

We can always write the left hand side as

$$\rho(\mathcal{H}(\boldsymbol{\Gamma})) = \int d\boldsymbol{\Gamma}' \rho(\mathcal{H}(\boldsymbol{\Gamma}'))\delta(\boldsymbol{\Gamma}'-\boldsymbol{\Gamma}), \tag{24}$$

so we have a functional of ρ which is identically zero,

$$\int d\boldsymbol{\Gamma}' \rho(\mathcal{H}(\boldsymbol{\Gamma}'))\left[\delta(\boldsymbol{\Gamma}'-\boldsymbol{\Gamma}) - \frac{\exp\left(-B(\boldsymbol{\Gamma}')\mathcal{H}(\boldsymbol{\Gamma})\right)}{Z(B(\boldsymbol{\Gamma}'))}\right] = 0. \tag{25}$$

Now we will take the functional derivative with respect to ρ on both sides and assume that B is independent of ρ, that is, $\delta B/\delta\rho(E) = 0$. It follows that

$$\frac{\exp\left(-B(\boldsymbol{\Gamma}')\mathcal{H}(\boldsymbol{\Gamma})\right)}{Z(B(\boldsymbol{\Gamma}'))} = \delta(\boldsymbol{\Gamma}'-\boldsymbol{\Gamma}). \tag{26}$$

Integrating with respect to $\boldsymbol{\Gamma}'$ we get

$$\int d\boldsymbol{\Gamma}' \exp\left(-B(\boldsymbol{\Gamma}')\mathcal{H}(\boldsymbol{\Gamma})\right) = Z(B(\boldsymbol{\Gamma})) \tag{27}$$

therefore, $B(\boldsymbol{\Gamma})$ depends on $\boldsymbol{\Gamma}$ only through $\mathcal{H}(\boldsymbol{\Gamma})$. Using this, we can write Eq. (23) as

$$\rho(E) = \int_0^\infty dE'\Omega(E')\rho(E')\frac{\exp\left(-B(E')E\right)}{Z(B(E'))}, \tag{28}$$

which again, can be rewritten as

$$\int_0^\infty dE'\rho(E')\left[\delta(E'-E) - \Omega(E')\frac{\exp\left(-B(E')E\right)}{Z(B(E'))}\right] = 0. \tag{29}$$

As this must be valid for any ρ, we take the functional derivative $\delta/\delta\rho$ and assume B does not depend on ρ. It follows that

$$\Omega(E')\frac{\exp\left(-B(E')E\right)}{Z(B(E'))} = \delta(E'-E), \tag{30}$$

for any pair of values E and E'. This, however, cannot be fulfilled by any function $B(E)$, as we will show in what follows.

First, it may seem obvious by simple inspection that the left-hand side is in general a positive function of E and E', not necessarily a delta function. However, note that as a function of E' it can still become as sharply-peaked as needed around E if the density of states Ω grows fast enough and the exponential factor falls fast enough. To prove that this is not the case, imagine fixing $E' = E_0$ so that $0 < B(E_0) < \infty$. Let us call $\beta_0 = B(E_0)$ and $Q = \Omega(E_0)/Z(\beta_0)$. Then we have

$$Q\exp\left(-\beta_0 E\right) = \delta(E_0 - E), \tag{31}$$

for all possible values of E. Now, choosing $E = E_0 \pm \Delta E$ with $0 < |\Delta E| < E_0$, we see from Eq. (31) that

$$\exp(-\beta_0\Delta E) = \exp(\beta_0\Delta E) = 0, \tag{32}$$

which is a contradiction for finite values of β_0 and $|\Delta E|$. This proves the theorem.

Despite this proof of impossibility we can provide a useful definition of inverse temperature, namely

$$\beta_{\mathcal{S}} := \left\langle\beta\right\rangle_{\mathcal{S}}, \tag{33}$$

as the **expectation of the parameter** β in the state $\mathcal{S}$. This allows us to define the temperature as simply $k_B T_{\mathcal{S}} = 1/\beta_{\mathcal{S}}$. The inverse temperature $\beta_{\mathcal{S}}$ can be computed from estimators $\hat{\beta}(\boldsymbol{r},\boldsymbol{p})$ and this is a value one can use to compare different states or to approximate the ensemble to the nearest canonical ensemble.

In order to show the validity of temperature estimators in an ensemble $P(\boldsymbol{\Gamma}|\mathcal{S})$ such as the one in Eq. (10) (of which superstatistics is a particular case), we make use of the conjugate variables theorem (CVT) [21] for the canonical ensemble (a brief proof of which is given in the Appendix),

$$\left\langle\nabla\cdot\boldsymbol{v}\right\rangle_\beta = \beta\left\langle\boldsymbol{v}\cdot\nabla\mathcal{H}\right\rangle_\beta \tag{34}$$

and marginalize over β, using the identity

$$\Big\langle g(\beta,\mathbf{\Gamma})\Big\rangle_{\mathcal{S}} = \int_0^{\infty} d\beta P(\beta|\mathcal{S})\Big\langle g(\beta,\mathbf{\Gamma})\Big\rangle_{\beta}. \tag{35}$$

We see that for the state $\mathcal{S}$ the following form of the CVT holds,

$$\Big\langle \nabla\cdot\boldsymbol{v}\Big\rangle_{\mathcal{S}} = \Big\langle \beta\boldsymbol{v}\cdot\nabla\mathcal{H}\Big\rangle_{\mathcal{S}}, \tag{36}$$

in which β is taken as an additional degree of freedom, and the expectation is taken under the joint distribution $P(\mathbf{\Gamma},\beta|\mathcal{S})$. Choosing

$$\boldsymbol{v} = \frac{\boldsymbol{\omega}}{\boldsymbol{\omega}\cdot\nabla\mathcal{H}} \tag{37}$$

as in Ref. [21], we find that

$$\beta_{\mathcal{S}} = \Big\langle \hat{\beta}\Big\rangle_{\mathcal{S}} = \Big\langle \nabla\cdot\Big[\frac{\boldsymbol{\omega}}{\boldsymbol{\omega}\cdot\nabla\mathcal{H}}\Big]\Big\rangle_{\mathcal{S}}. \tag{38}$$

for any $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{r},\boldsymbol{p})$. It seems suggestive to associate $\hat{\beta}$ with B but the point of our proof is that precisely, this choice (or any other) cannot reproduce all the moments of $P(\beta|\mathcal{S})$.

For the particular case of $\boldsymbol{\omega} = \boldsymbol{p}/m$, we obtain a kinetic expression

$$\beta_{\mathcal{S}} = \frac{1}{k_B T_{\mathcal{S}}} = \frac{3N-2}{2}\Big\langle K^{-1}\Big\rangle_{\mathcal{S}} \tag{39}$$

with K the kinetic energy of the system. Note that, because $\big\langle K^{-1}\big\rangle > \big\langle K\big\rangle^{-1}$ by Jensen's inequality [22],

$$T_{\mathcal{S}} < \frac{2}{(3N-2)k_B}\Big\langle K\Big\rangle_{\mathcal{S}}. \tag{40}$$

and so the intuitive generalization of Eq. (1) overestimates the temperature.

6. Conclusions

The theorem just proven rules out any intrinsic definition of temperature as a phase-space function in superstatistics. In statistical terms, we can say that the probability distribution $P(\beta|\mathcal{S})$ is not a sampling distribution, and β has to be interpreted as a parameter.

Our findings do not diminish the power of the superstatistical formalism or attempt to undermine its foundations. On the contrary, we are led to the conclusion that the notion of instantaneous or local temperature is at fault and that it might be separated from the idea of pure superstatistics, where β is kept as a parameter. There are already efforts to conceptually reformulate superstatistics from a Bayesian point of view [15], in which one does not need actual variations (temporal or spatial) of a physical quantity. Instead there are uncertainties in the well-defined and unique (but unknown) value of β.

Acknowledgments

SD gratefully acknowledges funding from FONDECYT, Chile 1140514, and GG also acknowledges funding from FONDECYT, Chile 1171127.

Appendix. Simple proof of the conjugate variables theorem (CVT)

For an arbitrary distribution of microstates $P(\mathbf{\Gamma})$ let us construct the expectation of $\nabla\cdot\boldsymbol{\omega}(\mathbf{\Gamma})$, with $\boldsymbol{\omega}$ an arbitrary but differentiable vector field,

$$\Big\langle \nabla\cdot\boldsymbol{\omega}\Big\rangle = \int_V d\mathbf{\Gamma} P(\mathbf{\Gamma})(\nabla\cdot\boldsymbol{\omega}). \tag{41}$$

We consider the divergence theorem applied to a volume V with boundary Σ and $\mathbf{v} = \boldsymbol{\omega}(\mathbf{\Gamma})P(\mathbf{\Gamma})$,

$$\int_V d\mathbf{\Gamma}(\nabla\cdot\mathbf{v}) = \int_{\Sigma} d\mathbf{\Sigma}\cdot\mathbf{v}. \tag{42}$$

We obtain

$$\begin{aligned}\int_V d\mathbf{\Gamma}\Big[P(\mathbf{\Gamma})\nabla\cdot\boldsymbol{\omega} + \omega\cdot\nabla P(\mathbf{\Gamma})\Big] &= \int_{\Sigma} d\mathbf{\Sigma}\cdot\boldsymbol{\omega}(\mathbf{\Gamma})P(\mathbf{\Gamma}) \\ &= 0,\end{aligned} \tag{43}$$

if the probability P is zero on the boundary Σ. [23]

By replacing ∇P as $P\nabla \ln P$ we can write both integrals in the left hand side as expectations over P, and finally obtain the CVT in its general form,

$$\Big\langle \nabla \cdot \boldsymbol{\omega}(\boldsymbol{\Gamma})\Big\rangle + \Big\langle \boldsymbol{\omega}(\boldsymbol{\Gamma}) \cdot \nabla \ln P(\boldsymbol{\Gamma})\Big\rangle = 0. \tag{44}$$

Replacing $P(\boldsymbol{\Gamma})$ by $P(\boldsymbol{\Gamma}|\beta)$ in Eq. (6) we get the canonical version of CVT, Eq. (34).

References

[1] C. Beck, E.G.D. Cohen, Physica A 322 (2003) 267.
[2] C. Beck, Contin. Mech. Thermodyn. 16 (2004) 293.
[3] A.M. Reynolds, Phys. Rev. Lett. 91 (2003) 084503.
[4] P.H. Chavanis, Physica A 359 (2006) 177.
[5] A. Porporato, G. Vico, P.A. Fay, Geophys. Res. Lett. 33 (2006) L15402.
[6] R. Hanel, S. Thurner, M. Gell-Mann, Proc. Nac. Acad. Sci. 108 (2011) 6390.
[7] S. Thurner, B. Corominas-Murtra, R. Hanel, Phys. Rev. E 96 (2017) 032124.
[8] C. Kittel, Phys. Today 41 (1988) 93.
[9] B.B. Mandelbrot, Phys. Today 42 (1989) 71.
[10] G.S. Boltachev, J.W.P. Schmelzer, J. Chem. Phys. 133 (2010) 134509.
[11] M. Falcioni, D. Villamaina, A. Vulpiani, A. Puglisi, A. Sarracino, Amer. J. Phys. 79 (2011) 777.
[12] P.D. Dixit, J. Chem. Phys. 138 (2013) 184111.
[13] P.D. Dixit, Phys. Chem. Chem. Phys. 17 (2015) 13000.
[14] C. Tsallis, J. Stat. Phys. 52 (1988) 479.
[15] F. Sattin, Eur. Phys. J. B 49 (2006) 219.
[16] H.H. Rugh, Phys. Rev. Lett. 78 (1997) 772.
[17] G. Rickayzen, J.G. Powles, J. Chem. Phys. 114 (2001) 4333.
[18] G. Palma, G. Gutiérrez, S. Davis, Phys. Rev. E 94 (2016) 62113.
[19] C. Beck, Phil. Trans. R. Soc. A 369 (2011) 453.
[20] M.E. Tuckerman, Statistical Mechanics: Theory and Molecular Simulation, Oxford University Press, 2010.
[21] S. Davis, G. Gutiérrez, Phys. Rev. E 86 (2012) 051136.
[22] T.M. Cover, J.A. Thomas, Elements of Information Theory, John Wiley and Sons, 2006.
[23] If this is not the case, it is always possible to augment the region V to V' and redefine the probability P, so that V' includes the original integration volume V and assigns zero probability to states outside V, for instance via a Heaviside step function.