

Meaning of temperature in different thermostistical ensembles

Peter Hänggi,^{1,2} Stefan Hilbert,^{3,*} and Jörn Dunkel⁴

¹*Institute of Physics, University of Augsburg, Universitätsstraße 1, D-86135 Augsburg, Germany*

²*Nanosystems Initiative Munich, Schellingstr. 4, D-80799 München, Germany*

³*Exzellenzcluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany*

⁴*Department of Mathematics, Massachusetts Institute of Technology,
77 Massachusetts Avenue E17-412, Cambridge, MA 02139-4307, USA*

(Dated: July 22, 2015)

Depending on the exact experimental conditions, the thermodynamic properties of physical systems can be related to one or more thermostistical ensembles. Here, we survey the notion of thermodynamic temperature in different statistical ensembles, focusing in particular on subtleties that arise when ensembles become non-equivalent. The ‘mother’ of all ensembles, the microcanonical ensemble, uses entropy and internal energy (the most fundamental, dynamically conserved quantity) to derive temperature as a secondary thermodynamic variable. Over the past century, some confusion has been caused by the fact that several competing microcanonical entropy definitions are used in the literature, most commonly the volume and surface entropies introduced by Gibbs. It can be proved, however, that only the volume entropy satisfies exactly the traditional form of the laws of thermodynamics for a broad class of physical systems, including all standard classical Hamiltonian systems, regardless of their size. This mathematically rigorous fact implies that negative ‘absolute’ temperatures and Carnot efficiencies > 1 are not achievable within a standard thermodynamical framework. As an important offspring of microcanonical thermostistics, we shall briefly consider the canonical ensemble and comment on the validity of the Boltzmann weight factor. We conclude by addressing open mathematical problems that arise for systems with discrete energy spectrum.

PACS numbers: 05.20.-y, 05.30.-d, 05.70.-a

Keywords: Thermodynamic ensembles, entropy, isolated systems, weak and strong coupling

I. INTRODUCTION

The fundamental differential relation [1]

$$\frac{1}{T} = \frac{\partial S}{\partial E} \quad (1)$$

connects the thermodynamic state functions temperature T , internal energy E , and entropy S . Given S as a function of E , and possibly other control parameters, Eq. (1) in fact *defines* the thermodynamic temperature. The concept of entropy was introduced by Rudolf Clausius [2] in 1865. Clausius chose the symbol S in honor of (Nicolas Léonard) Sadi Carnot, who laid the groundwork for the Second Law of Thermodynamics. The celebrated Clausius relation $dS = \delta Q/T$ identifies the inverse of the thermodynamic temperature T as the integrating factor for the Second Law, with $\delta Q \stackrel{\cong}{\leq} 0$ denoting quasi-static and reversible infinitesimal heat exchange. After Clausius’ seminal paper [2], it took about 30 more years until Gibbs [3], Einstein, Planck [4, 5] and others [6] were able to connect firmly thermodynamics and statistical mechanics – and yet certain aspects of this connection have remained a subject of debate up to this day.

The standard approach in statistical mechanics is to identify thermodynamic state functions with specific average values of a suitably chosen statistical ensemble that correctly reflects the physical conditions under which measurements are performed (perfect isolation, coupling to an energy or matter reservoir, etc.). The most fundamental statistical ensemble is the microcanonical ensemble (MCE), describing the thermodynamics of isolated systems that are governed by energy conservation and which, at equilibrium, cannot exchange heat or matter with their surroundings. The MCE is the foundation of other thermostistical ensembles, including the canonical ensemble (which permits permanent energy and/or heat exchange with the environment) and the grand-canonical ensemble (which allows both energy and matter exchange). These two subordinate ensembles can be derived from the MCE by considering a subsystem of interest that is *weakly* coupled to the rest of the globally isolated microcanonical system, which is then interpreted as an environment (heat bath or particle reservoir) for the particular subsystem.

*Electronic address: stefan.hilbert@tum.de

Recent experimental advances make it possible to investigate thermodynamic properties of very small systems (single molecules, Brownian colloids or even individual atoms) that may be, in good approximation, decoupled from the environment or that can be in weak or strong contact with a much larger system. Such finite-system studies provide a valuable testbed for the notion and meaning of thermodynamic temperature in the context of various statistical ensembles. Particularly interesting from a theoretical and practical perspective are situations in which different ensemble descriptions are not guaranteed to be equivalent. Ensemble inequivalence is more norm than exception in finite-size systems but can also occur in macroscopic systems with long-range interactions or when the density of states (DoS) is a nonmonotonic function of energy. Equilibrium systems of the latter type are often classified as anomalous [1] and, if entropy is chosen naively, they can give rise to the paradoxical notion of a negative ‘absolute’ temperature.

In the remainder of this contribution, we will survey the meaning of temperature in thermodynamics by summarizing and commenting on results from recent more detailed studies [7, 8].

II. MICROCANONICAL THERMODYNAMICS AND ABSOLUTE TEMPERATURE

The MCE describes the thermostatics of a strictly isolated system through the density operator $\rho = \delta(E - H)/\omega$, where the normalization constant ω is the density of states (DoS). The MCE is the most fundamental ensemble as it only relies on the conservation of energy E , arising from the time-translation invariance of the underlying Hamiltonian H . External thermodynamic control parameters Z , such as available system volume, particle numbers, electric or magnetic fields, enter as parameters through the Hamiltonian $H(Z)$ and the DoS $\omega(E, Z)$. To connect the MCE to thermodynamics, J. W. Gibbs [3] studied two different candidates for the thermodynamic entropy of an isolated system. The first is the *volume entropy*, which in modern notation takes the form

$$S_G = k_B \ln \Omega(E, Z). \quad (2)$$

Here, k_B denotes the Boltzmann constant and the dimensionless *volume*-quantity $\Omega(E, Z)$ is the integrated DoS, classically obtained by integrating the non-negative DoS $\omega \geq 0$ up to energy E ,

$$\Omega(E, Z) = \int_0^E dE' \omega(E', Z), \quad (3)$$

assuming zero ground-state energy for a physically stable system. Since Ω is non-decreasing function of E , the temperature T_G obtained from S_G and Eq. (1) is strictly non-negative.

For classical Hamiltonian systems $H(\xi, Z)$ with phase-space variables ξ , the integrated DoS $\Omega(E, Z)$ equals the properly normalized (via division by the symmetries of the degrees of freedom) and dimensionless (via division by the appropriate power of Planck’s constant) integrated phase space volume up to the energy E . We may write this formally as

$$\Omega(E, Z) = \mathcal{Tr}_\xi \Theta[E - H(\xi, Z)] \quad (4)$$

where Θ denotes the unit-step function and \mathcal{Tr} the phase-space integral over distinguishable microstates ξ . For isolated quantum systems with discrete energy spectrum, ξ comprises the complete set of quantum numbers, and we may interpret Ω in Eq. (4) as a discrete level counting function, defined on the spectrum $\{E_n\}$ of the Hamiltonian. Intuitively, the discrete function $\Omega(E_n, Z)$ sums the eigenspace dimensions of the eigenvalues $E_j \leq E_n$. In the quantum case, one needs to postulate additional smoothing procedures before one can apply differential thermodynamic relations such as Eq. (1) (see discussion in Sec. IV below).

Following Gibbs’ seminal work, Hertz [6] demonstrated the mechanical adiabatic invariance of the volume entropy S_G for classical systems. The exact connection between S_G , its corresponding temperature T_G and equipartition for classical finite size systems was emphasized in early works by Schlüter [9] and Khinchin [10]. More recent discussions and applications of the Gibbs’ volume entropy can be found in Refs. [7, 11–17].

The second microcanonical entropy candidate studied by Gibbs [3] is the *surface entropy*

$$S_B = k_B \ln[\omega(E, Z)\epsilon]. \quad (5)$$

The quantity ϵ denotes an arbitrary energy constant, needed to make the argument of the logarithm dimensionless. That the definition of S_B requires such an additional *ad hoc* parameter is conceptually unappealing, but bears no relevance for thermodynamic quantities that are related to derivatives of $S_B(E, Z)$ – provided ϵ is assumed to be independent of (E, Z) . One can show however that the presence of ϵ can cause S_B to violate Planck’s formulation of

the Second Law [7]. For discrete quantum systems with singular DoS ω , Eq. (5) also requires additional interpolation and/or smoothing procedures (see Sec. IV). The subscript ‘B’ in Eq. (5) signals that this definition is also commonly referred to as Boltzmann entropy nowadays, which unfortunately does not seem to reflect properly the actual history. Although Boltzmann’s tombstone famously carries the entropy formula $S = k_B \log W$, it was, according to Arnold Sommerfeld [18], Max Planck [4, 5] who first established this relation. As described in many textbooks, the entropy expression S_B in Eq. (5) can be heuristically obtained by identifying $\log = \ln$ and interpreting $W = \epsilon\omega(E, Z)$ as the number of microstates accessible to a physical system at energy E . This may explain the popularity of the term ‘Boltzmann entropy’.

It is well known that for macroscopic *normal* systems [1] with a large number of microscopic degrees of freedom, most of the phase space volume is contained in a narrow shell just below the energy E . In such cases, the two entropy definitions become essentially indistinguishable and predict practically identical thermodynamic equations of state. There exists, however, a wide range of systems for which S_B and S_G are non-equivalent.

A. Self-consistency checks

The question as to whether S_G or S_B are viable candidates for the thermodynamic entropy of isolated systems, can be answered directly by testing either candidate against the Laws of Thermodynamics. The approximation-free analysis in Ref. [7] shows that for a broad class of physical systems, which includes all standard classical Hamiltonian systems¹ of *arbitrary* size, the Gibbs volume entropy S_G satisfies the traditional formulations of the Zeroth, First and Second Law exactly. In contrast, the surface entropy S_B is found to violate these laws in many situations [3, 7, 8]. While referring the reader to Ref. [7] for technical details, we briefly summarize the most essential results.

i. That S_G , but not S_B , satisfies the Zeroth Law is a reflection of the fact [7, 9, 10, 12, 13, 17] that only S_G satisfies the microcanonical equipartition theorem exactly. More precisely, denoting the microcanonical averages by $\langle \cdot \rangle_E$, the Stokes theorem implies [10] implies that, for all standard classical Hamiltonian systems with topologically simple phase space \mathbb{R}^d , the equipartition identity²

$$k_B T_G = \left(\frac{\partial S_G}{\partial E} \right)^{-1} = \left\langle \xi_k \frac{\partial H}{\partial \xi_k} \right\rangle \quad (6)$$

holds for *any* of the canonical coordinates (ξ_1, \dots, ξ_d) . By contrast, this relation is *in general* violated for the surface entropy, ruling out S_B as a consistent thermodynamic entropy.

ii. Compliance of $S = S_G$ and $T = T_G$ with the First Law

$$dS = \frac{1}{T} dE + \sum_i \frac{F_i}{T} dZ_i, \quad F_i := T \left(\frac{\partial S}{\partial Z_i} \right)_E \stackrel{!}{=} - \left\langle \frac{\partial H}{\partial Z_i} \right\rangle_E, \quad (7)$$

follows directly from a simple integration by parts [7, 16]. Note that the last equality in Eq. (7) ensures that statistical averages agree with thermodynamic observables. One can easily verify that this consistency criterion is, in general, violated for the Boltzmann entropy S_B .

iii. Planck’s Second Law of Thermodynamics for isolated systems can be, in essence, stated as follows: Consider two isolated microcanonical systems that are initially separated and have entropies $S_1(E_1)$ and $S_2(E_2)$, respectively. Now couple the two systems weakly and let them equilibrate. Assuming energy conservation throughout the process, the joint equilibrated system is again microcanonical and has entropy $S_{1+2}(E_{1+2}) = S_{1+2}(E_1 + E_2)$. Then, Planck’s Second Law demands that the entropy of the final state is larger than the sum of the initial entropies,

$$S_{1+2}(E_{1+2}) \geq S_1(E_1) + S_2(E_2). \quad (8)$$

¹ These are confined systems with quadratic kinetic energy and finite ground-state energy.

² For systems with complex phase space topology, Eq. (6) can be violated, see example in Sec. IIB below, where the phase space regions corresponding clockwise and anti-clockwise motion become disconnected for supercritical energy values. Such topologically peculiar systems do not thermalize in the traditional sense. However, for the most commonly considered standard classical Hamiltonian systems, Eq. (6) is strictly satisfied.

Basic integral convolution properties imply that Eq. (8) is always satisfied for S_G (in most cases, even with strictly ' $>$ ') but not necessarily by S_B [7].

In this context, it is worthwhile to note that any subsequent attempt to decouple the two systems results in non-microcanonical distributions for the separated systems, since the exact individual energies are not known anymore due to the permanent energy exchange during the equilibration phase (i.e., thermal coupling is irreversible). This means that, without further manipulation or measurements (or the introduction of a Maxwell demon), the total entropy remains $S_{1+2}(E_1 + E_2)$ after separation, a fact that has been missed by authors [19] who recently criticized the Gibbs entropy. Unsurprisingly, this basic error led to paradoxical conclusions [19], such as an apparent violation of mathematically exact inequalities.

For completeness, we mention that previous studies rarely focused on the Third Law, mainly because it is well known that many classical systems (including the ideal gas) do not obey the Third Law. Typically, verification of the Third Law requires a consistent quantum-mechanical treatment³. Evidently, the Gibbs entropy satisfies $S_G(E_0) = k_B \ln g_0$ with g_0 denoting the degeneracy of the ground state energy E_0 and hence fulfills the most basic version of the Third Law.

B. Positive and negative absolute temperatures: an example

The primary thermodynamic state variables of an isolated system with Hamiltonian $H(Z) = E$ are energy E and control parameters Z . By contrast, the temperature T is a secondary derived quantity determined by Eq. (1). For the Gibbs volume entropy, one finds explicitly

$$k_B T_G = \frac{\Omega(E, Z)}{\omega(E, Z)}, \quad \omega(E, Z) = \frac{\partial \Omega}{\partial E}. \quad (9)$$

Since both the integrated DoS $\Omega \geq 0$ and the DoS $\omega \geq 0$ are non-negative, the Gibbs temperature is strictly non-negative. For comparison, the Boltzmann temperature is given by

$$k_B T_B = \frac{\omega(E, Z)}{\nu(E, Z)}, \quad \nu(E, Z) = \frac{\partial \omega}{\partial E}. \quad (10)$$

The Boltzmann temperature T_B is negative whenever the DoS ω is a locally decreasing with E (see example in Fig. 1). This happens for Ising-type spin or laser systems in the population inverted system, as well is in Hamiltonian systems exhibiting singular points in their DoS that separate regions of $\nu(E, Z) > 0$ with regions with negative-valued $\nu(E, Z)$.

An instructive ergodic 1D example [17] is the classical pendulum (mass m , length L , gravitational acceleration g) with Hamiltonian

$$H(\phi, p_\phi) = \frac{p_\phi^2}{2m} - mgL \cos \phi. \quad (11)$$

The energy range is bounded from below but unbounded from above, $0 < E < \infty$. The DoS ω can be given analytically in terms of complete elliptic integrals of the first kind (see in Ref. [17] and Fig. 1 therein). The DoS increases in the oscillatory regime $0 < E < mgL$, exhibits a singularity at $E = mgL$, where the orbit period diverges, and decreases in the continuous-rotation regime, $mgL < E < \infty$. Accordingly, $\nu(E > mgL) < 0$ decays monotonically towards zero as $E \rightarrow \infty$, implying a negative Boltzmann temperature for $E > mgL$. By contrast, the Gibbs temperature is positive for all $E > 0$. In particular, for $E \gg mgL$, any further increase of energy is, in essence, purely kinetic and the system approaches an ideal one-particle gas on a circle, which should asymptotically satisfy $E = +\frac{1}{2}k_B T$, unless one is willing to give up this standard caloric equation of state. It easy to check that this relation holds only for $T = T_G$.

C. Additional remarks

Heat does not always flow from hot to cold. The thermodynamic state of an isolated system is completely determined by the primary state variables (E, Z) . Since temperature is not a primary state variable, it cannot, in

³ As a note of caution: One can find many partially conflicting versions of the Third Law in the literature, and some naive formulations are not applicable to isolated systems, or only apply to systems with non-degenerate ground-state or finite energy gap between ground-state and lowest excited energy levels.

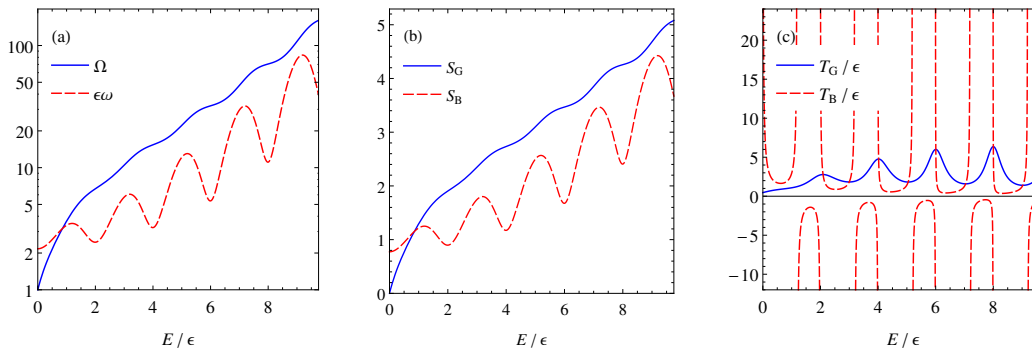


FIG. 1: Non-uniqueness of microcanonical temperatures for a system with nonmonotonic density of states (DoS); figure adapted from Ref. [7]. (a): DoS ω (red, dashed) and integrated DoS Ω (blue) for the example in Eq. (23) of Ref. [7]. (b): Gibbs volume entropy S_G (blue) and Boltzmann surface entropy S_B (red, dashed) in units $k_B = 1$. (c): Gibbs temperature T_G (blue) and Boltzmann temperature T_B (red, dashed). The example illustrates that, in general, neither the Gibbs nor the Boltzmann temperature uniquely characterize the thermal state of an isolated system because the same temperature value can correspond to different energy values.

general, uniquely characterize the thermodynamic state of a microcanonical system (see Fig. 1). This means there exist situations where neither the Gibbs temperature T_G nor the Boltzmann temperature T_B can predict the energy flow between weakly coupled systems that had different temperatures before contact [7]. For instance, for a system with oscillatory DoS, two or more significantly different energy values can have the same temperature, regardless of which entropy definition one adopts (see Fig. 1c). Therefore, the naive formulation of the Second Law ‘heat always flows from hot to cold’ does not hold in general. Likewise invalid are versions of the Zeroth Law that claim that isolated systems with equal temperatures should not produce a net heat flow between them when brought into thermal contact. This can again be readily seen by considering, for example, the coupling of an ideal gas to a system with oscillatory DoS as in Fig. 1.

Thermodynamics applies to equilibrium systems of any size. It takes little effort to verify that Eqs. (6)-(8) hold exactly for standard classical Hamiltonian systems with an arbitrary number of degrees of freedom N . Similarly, the canonical ensemble discussed below can be applied to (sub-)systems of any size. These mathematical facts are widely appreciated by many colleagues [8, 11, 14, 15, 17] – in particular those interested in understanding DNA folding [21], microscopic information storage and erasure [22] and fluctuation phenomena [23] – and yet remain ignored by others [19, 24]. When judged objectively, there is no doubt that the application of thermodynamic concepts to finite systems has considerably advanced our understanding of biophysical, colloidal and quantum processes⁴. Compared with infinite-systems thermodynamics, a practical difference is given by the fact that fluctuations generally play a (much) more important role in small systems. The presence of fluctuations, of course, does not mean that it is forbidden to characterize single DNA molecules thermodynamically; on the contrary, such fluctuations typically contain important additional thermodynamic and energetic information that is usually lost in the infinite-system limit. Therefore, it would seem wiser to focus on understanding better the fluctuations of thermostatical variables in finite systems, such as those of virial quantities on the rhs. of Eq. (7), instead of discarding finite-system thermodynamics on purely habitual grounds [24]. Dogmatic insistence on the thermodynamic limit $N \rightarrow \infty$ is about as useful as insisting on the Newtonian limit, corresponding to speed of light $c \rightarrow \infty$, in relativity. In both cases, things may become simpler, but one is missing out on relevant physics.

Clausius relation & Carnot efficiency. Campisi [8] showed recently that the inverse Gibbs temperature T_G^{-1} appears naturally as the integrating factor in the Clausius relation for virtually all practically relevant physical systems. This corroborates the fact that the Gibbs temperature T_G should be identified with the *absolute* thermodynamic temperature T , unless one is willing to abandon the Clausius relation. Moreover, the non-negativity of the Gibbs temperature directly implies that Carnot efficiencies cannot exceed 1.

⁴ Most of the experimental applications involve the canonical ensemble, as DNA molecules [21] or colloids [22, 23] are typically held in a liquid bath that acts as a canonical thermostat. However, if one accepts the applicability of canonical thermostatics to finite systems, then there exist no mathematical or logical or physical reasons that forbid the application of microcanonical statistics to finite systems, because Eqs. (6)-(8) hold for systems of any size. Of course, the thermostatical characterization of small systems should not just be limited to the mean values appearing in Eqs. (6) and (7) but should also include a careful fluctuation analysis of the underlying stochastic observables.

Thermodynamic potentials can be ‘nonlocal’. It is sometimes argued that the Gibbs entropy cannot be the ‘correct’ thermodynamic entropy as it is based on the integrated phase-space volume Ω , which is a ‘nonlocal’ quantity that arises from a summation over states in an extended energy range. This argument might appear superficially appealing but it is ill-founded for (at least) two reasons: First, the microcanonical averages appearing in Eqs. (6) and (7) are computed purely locally on the energy surface in phase space. Yet, the Stokes theorem implies that they can be related to the enclosed phase-space volume [10] and, hence, entropy should be a nonlocal volume-related quantity. Second, if we insisted on purely local potentials everywhere in physics then we would have to stop using force potentials in mechanics, which are in essence ‘non-local’ integrals over local forces experienced by particles. At this point, however, it is helpful to recall why such non-local potentials are introduced in mechanics in the first place: They allow us to define an important conserved quantity, *energy*. Just as the energy is invariant under infinitesimal time translations, the integrated phase-space volume Ω is invariant under infinitesimal adiabatic parameter translations [6]. Hence, it should not be surprising but rather be expected that thermodynamic potentials may be non-local in energy space.

Ising models are bad benchmarks. When using specific theoretical models to illustrate alleged pros and cons of certain entropy definitions, then it is advisable to verify first that these models respect superordinate experimentally established knowledge. Specifically, while the observed stability of matter implies the existence of lower energy bounds on Hamiltonians, there exists no evidence to date for strict upper energy bounds. This means that $E \rightarrow -E$ is not a fundamental symmetry of physics and, hence, one should not impose such energy-reflection symmetry on thermodynamic quantities. For the same reason, it is not advisable to base arguments exclusively on Ising-type models, which are *ad hoc* truncations of more fundamental Hamiltonians that are *not* bounded from above, if one wants to evaluate the conceptual validity of a certain thermostistical framework.

Ensemble (in)equivalence. Although S_G, S_B and other entropy candidates [7] often yield practically indistinguishable predictions for the thermodynamic properties of large *normal* systems [1], such as quasi-ideal gases with macroscopic particle numbers, they can produce substantially different predictions for finite mesoscopic systems, for ad-hoc truncated Hamiltonians with upper energy bounds or even for macroscopic gravitational systems [25]. This implies (see discussion in Sec. III below) that, in general, microcanonical and canonical descriptions are not equivalent, which is not surprising as they refer to different physical conditions (complete isolation *vs.* coupling to an infinite bath).

III. TEMPERATURE IN THE CANONICAL ENSEMBLE

The canonical Boltzmann factor $e^{-\beta H}$, where $T = (k_B \beta)^{-1}$ is commonly identified with the bath temperature, has become one of the most frequently employed statistical tools in physics. It is therefore conceptually important and practically useful to understand potential validity limits, which arise from assumptions and approximations made during the derivation from the underlying MCE.

A. Boltzmann factor and temperature

We briefly summarize the key assumptions underlying the derivation of the Boltzmann factor by considering a system of interest \mathcal{S} which is coupled to another system \mathcal{B} that acts as a heat bath. The starting point of the derivation is the microcanonical density operator $\rho^{\mathcal{J}}(\xi|E^{\mathcal{J}}, Z) = \delta[E^{\mathcal{J}} - H^{\mathcal{J}}(\xi, Z)]/\omega^{\mathcal{J}}(E^{\mathcal{J}}, Z)$ of the total system $\mathcal{J} = \mathcal{S} + \mathcal{B}$. The first key assumption en route to the Boltzmann factor is *weak coupling*, which means that one neglects system-bath interaction contributions to the total energy and Hamiltonian, by writing $E^{\mathcal{J}} = E^{\mathcal{S}} + E^{\mathcal{B}}$ and $H^{\mathcal{J}} = H^{\mathcal{S}} + H^{\mathcal{B}}$. Because the total energy $E^{\mathcal{J}}$ is fixed, the probability weight $P(E^{\mathcal{S}}|E^{\mathcal{J}}, Z)$ to find the fluctuating energy value $E^{\mathcal{S}}$ of the subsystem \mathcal{S} is simply given by [1, 7, 26]

$$P(E^{\mathcal{S}}|E^{\mathcal{J}}, Z) = \frac{\omega^{\mathcal{S}}(E^{\mathcal{S}})\omega^{\mathcal{B}}(E^{\mathcal{J}} - E^{\mathcal{S}})}{\omega^{\mathcal{J}}(E^{\mathcal{J}})} \quad (12)$$

Here, $\omega^{\mathcal{S}}(E^{\mathcal{S}})$ denotes the degeneracy of the subsystem energy value $E^{\mathcal{S}}$, while $\omega^{\mathcal{B}}(E^{\mathcal{B}})$ is the DoS of the bath at energy $E^{\mathcal{B}} = E^{\mathcal{J}} - E^{\mathcal{S}}$, and $\omega^{\mathcal{J}}(E^{\mathcal{J}})$ the total DoS at the total energy $E^{\mathcal{J}}$. Equation (12) can be equivalently rewritten as

$$P(E^{\mathcal{S}}|E^{\mathcal{J}}, Z) = \frac{\omega^{\mathcal{S}}(E^{\mathcal{S}})}{\epsilon \omega^{\mathcal{J}}(E^{\mathcal{J}})} \exp \left[\frac{S_{\mathcal{B}}^{\mathcal{B}}(E^{\mathcal{J}} - E^{\mathcal{S}})}{k_B} \right] \quad (13)$$

where $S_B^{\mathcal{B}}(E^{\mathcal{B}}) = S_B^{\mathcal{B}}(E^{\mathcal{J}} - E^{\mathcal{S}})$ denotes the Boltzmann entropy of the *bath*. As next step in the standard derivation [1, 26], one expands the Boltzmann entropy in the exponent around some conveniently chosen value $\bar{E}^{\mathcal{B}}$, typically taken to be the expectation value $E^{\mathcal{B}}$ of the bath energy⁵, keeping terms up to linear order:

$$S_B^{\mathcal{B}}(E^{\mathcal{J}} - E^{\mathcal{S}}) = S_B^{\mathcal{B}}(\bar{E}^{\mathcal{B}}) + \frac{1}{T_B^{\mathcal{B}}(\bar{E}^{\mathcal{B}})}(E^{\mathcal{J}} - E^{\mathcal{S}} - \bar{E}^{\mathcal{B}}) + \dots, \quad (14)$$

where $T_B^{\mathcal{B}} = (\partial S_B^{\mathcal{B}}/\partial E^{\mathcal{B}})^{-1}$ is Boltzmann temperature of the bath. Note that Eq. (14) is essentially an expansion in the energy fluctuations of the bath $\delta E^{\mathcal{B}} = E^{\mathcal{B}} - \bar{E}^{\mathcal{B}} = (E^{\mathcal{J}} - E^{\mathcal{S}}) - \bar{E}^{\mathcal{B}}$. Inserting the expansion (14) into Eq. (13) gives⁶

$$P(E^{\mathcal{S}}|E^{\mathcal{J}}, Z) = \frac{\omega^{\mathcal{S}}(E^{\mathcal{S}})}{\epsilon \omega^{\mathcal{J}}(E^{\mathcal{J}})} \exp \left[\frac{S_B^{\mathcal{B}}(\bar{E}^{\mathcal{B}})}{k_B} + \frac{(E^{\mathcal{J}} - \bar{E}^{\mathcal{B}}) - E^{\mathcal{S}}}{k_B T_B^{\mathcal{B}}(\bar{E}^{\mathcal{B}})} + \dots \right]. \quad (15)$$

Assuming *all* higher-order terms can be neglected, one obtains the standard result

$$P(E^{\mathcal{S}}|E^{\mathcal{J}}, Z) = \frac{\omega^{\mathcal{S}}(E^{\mathcal{S}})}{Z_C} \exp \left[-\frac{E^{\mathcal{S}}}{k_B T_B^{\mathcal{B}}(\bar{E}^{\mathcal{B}})} \right]. \quad (16)$$

where all remaining $E^{\mathcal{S}}$ -independent terms have been absorbed into the normalizing constant Z_C , which is the canonical partition function.

Thus, the temperature entering the celebrated Boltzmann factor $\exp(-\beta E^{\mathcal{S}})$ is the *Boltzmann temperature* $T_B^{\mathcal{B}}$ of the *bath*. One may therefore be tempted to assume that $T_B^{\mathcal{B}}$ can be identified with the thermodynamic temperature of the total system. However, this is logically incorrect because $T_B^{\mathcal{B}}(\bar{E}^{\mathcal{B}})$ is in general not equal [7] to the total system Boltzmann temperature $T_B^{\mathcal{J}}(E^{\mathcal{J}})$ or Gibbs temperature $T_G^{\mathcal{J}}(E^{\mathcal{J}})$, with the latter being the actual thermodynamic temperature. Of course, when the bath is macroscopically large and normal (e.g., ideal gas-like) from a thermostatical viewpoint, then the effective temperature $T_B^{\mathcal{B}}$ practically coincides with the Gibbs temperatures of the bath \mathcal{B} , the system \mathcal{S} , and the total system \mathcal{J} , and we have $T_B^{\mathcal{B}} = T_G^{\mathcal{B}} = T_G^{\mathcal{S}} = T_G^{\mathcal{J}}$ in this limit. In contrast, when considering finite thermostats (i.e., a bath with a finite number of microstates), then the exponential Boltzmann has to be replaced by a generalized Boltzmann factor, which may assume the form of a Tsallis-Renyi escort distribution [27].

B. Beyond weak coupling

Having surveyed the notion of temperature in the canonical ensemble, we still address another subtlety that concerns the heat capacities of nano-systems. When studying the thermodynamic properties of nano-scale devices, one naturally encounters the question whether the weak coupling approximation remains justified. Indeed, for a typical nano-system in contact with a heat bath, the coupling energy is usually of the order of the average system energy. Therefore, coupling terms in the Hamiltonian can no longer be neglected. This issue was investigated in detail in Refs. [28–30], which focused on the question how the canonical heat capacity is affected when nano-subsystems are *strongly* coupled to a large normal bath. These studies showed that the reduced canonical weight (or reduced density operator in quantum mechanics) of a strongly coupled small system is no longer of the Boltzmann form, when expressed in terms of the bare sub-system energy $E^{\mathcal{S}}$ or corresponding Hamilton operator $H^{\mathcal{S}}$. Instead, the canonical weight now features a renormalized subsystem Hamiltonian that depends explicitly on both effective bath temperature T_B and coupling strength. Moreover, as a main consequence, the canonical specific heat of the subsystem is not guaranteed to be positive and can, in fact, attain *negative* values even for $T_B > 0$ [28, 29]. The thermodynamic entropy⁷ of such a strongly coupled quantum system, obtained from its canonical partition function via the free energy, assumes a form that is close (but not exactly equal) to the quantum conditional entropy, and can become negative for $T_B > 0$ [30]. This does not affect, however, the validity of the Third Law as stated above, which holds true for $T_B \rightarrow 0$ even when a small quantum subsystem is strongly coupled to a heat bath; see Fig. 3 in Ref. [30] for an example.

⁵ Replacing the expansion point $\bar{E}^{\mathcal{B}}$ by the mode is not recommendable, as this procedure becomes ambiguous or even ill-defined when the DoS $\omega^{\mathcal{B}}(E^{\mathcal{B}})$ of the bath is oscillating or monotonically increasing.

⁶ The approximation (15) neglects all higher-order contributions in the Taylor expansion of $S_B^{\mathcal{B}}$. For example, the coefficient in front of the quadratic term is proportional to $\partial^2 S_B^{\mathcal{B}}/\partial^2 E^{\mathcal{B}} = -1/(T_B^2 C_B^{\mathcal{B}})$, where $C_B^{\mathcal{B}} = \partial E^{\mathcal{B}}/\partial T_B$ is the canonical specific heat of the bath. For this term to vanish individually, $C_B^{\mathcal{B}}$ has to be sufficiently large. Roughly speaking, one can expect that the second-order as well as higher-order expansion terms become negligible if the Boltzmann temperature $T_B^{\mathcal{B}}$ changes only slowly when the bath energy is varied. This typically requires a large bath.

⁷ For non-weakly coupled systems, correlations between bath and system cannot be neglected and, therefore, the thermodynamic entropy of such subsystems is no longer given by the classical Gibbs-Shannon entropy or the quantum-mechanical von-Neumann entropy.

C. Thermodynamic *vs.* information entropy

We conclude our discussion of the canonical ensemble with brief remarks on thermodynamic and information-theoretic entropies. The exponential Boltzmann distribution (16) is directly linked to the entropy $S_C = -k_B \mathcal{T}r(\rho \ln \rho)$, as already noted by Gibbs [3] who discussed S_C exclusively in the context of the canonical ensemble. Nowadays, S_C is commonly referred to as the *canonical* Gibbs-Shannon entropy in classical statistical mechanics and as the von Neumann entropy in quantum statistics. It is well-known that the canonical distribution (16) can be obtained by maximizing S_C under the assumption that the mean energy \bar{E}^S is given. However, such a purely formal ‘derivation’ conceals the underlying physical assumptions that determine the range of validity of the Boltzmann distribution (16). Also, entropy maximization arguments often leave the impression that there is a direct 1-to-1 correspondence between thermodynamics and information theory, which is somewhat misleading for a number of reasons: First, there exist many different information measures [31] and the Shannon entropy is just one of them - although an admittedly very nice one. Second, the Shannon entropy can be used to quantify the information content of arbitrary probability measures that, in most cases, have no relation to the thermodynamic equilibrium distributions. Third, the most fundamental equilibrium ensemble, the MCE, has a thermodynamic entropy that does not belong to the class of Shannon entropies. Therefore, some reservation seems in order when attempts are made to identify information-theoretic measures generically with thermodynamic entropies and *vice versa*. A similar note of caution applies when one tries to relate information-theoretic inequalities to thermodynamic inequalities that arise in the context of the Second Law or from thermodynamic stability considerations [26]. Potential analogies between thermodynamics and information theory are interesting and deserve to be explored in great detail, but they should not necessarily be raised to the level of postulates, when they have been shown to be incomplete and may obscure physical insight.

IV. OPEN QUESTIONS

The above discussion implicitly assumed that all derivatives exist and are well behaved. This is typically the case for classical Hamiltonian systems with the exception of critical points [13], as also encountered in the pendulum example above [17]. For quantum systems, the problem is generally more subtle since quantum-mechanical energy spectra can be partially or completely discrete, and are typically very sensitive to small perturbations that can break symmetry-related degeneracies. Similar effects occur in classical approximations to quantum systems, as for example the classical Ising model. Whenever one faces a completely or partially discrete spectrum $\{E_i\}$, the corresponding DoS $\omega(E, Z)$ becomes formally singular and essentially reduces to a collection of δ -function at those discrete energy values, $\omega(E) = \sum_i g_i \delta(E - E_i)$. In this case, the construction of a differentiable DoS requires some sort of smoothing procedure. This issue is closely related to the so-called ‘‘Weyl problem’’ of finding asymptotic approximations for the eigenvalue distributions of Hermitian operators in finite domains, by applying some suitable averaging procedure to obtain a continuous DoS [32]. For canonical systems, one typically uses such a smoothed DoS of the underlying energy spectrum at high ambient temperatures.

When the spectrum exhibits a discrete range, then one can define the integrated DoS $\Omega(E, Z)$, which enters the microcanonical Gibbs entropy S_G , in at least two different ways: The most commonly used method simply integrates the discrete DoS $\omega(E, Z)$ over E , which results in a step function $\tilde{\Omega}$ that gives rise to singular thermodynamic derivatives. This approach seems unsatisfactory mathematically, for one simply integrates over the ‘forbidden’ part⁸ of the spectrum while completely ignoring structural information encoded in the amplitude values g_i of ω in the interpolation regions (E_i, E_{i+1}) . A potentially better method [7] is based on analytic interpolation of the discrete level counting function $\Omega(E_n) = \sum_{E_j \leq E_n} \dim \mathcal{H}_j$, where \mathcal{H}_j is the eigenspace of E_j . Although the most natural interpolations appear obvious when $\Omega(E_n)$ can be written as $\Omega(E_n) = f(n)$ for some known function f (see examples in Ref. [7]), there remain open questions as to how to treat rigorously cases where no such closed-form representation is known.⁹

⁸ The union of intervals (E_i, E_{i+1}) .

⁹ Another practical, but not quite as elegant approach is to replace derivatives with finite differences, which in essence corresponds to linear and higher-order polynomial interpolations.

V. CONCLUSIONS

Gibbsian thermodynamics [3, 6] works consistently for finite and infinite systems, because the underlying mathematical and statistical foundations, most importantly Liouville’s theorem, merely rely on generic conservation laws that arise from the Hamiltonian structure of the microscopic dynamics. Working with infinite systems is generally easier as this limit often (but not always) forgives a certain laxness in defining entropy and thermostistical observables, since different definitions may show the same asymptotic behavior when the particle number N is let to ∞ . As mentioned before, this is quite analogous to the fact that the Newtonian limit $c \rightarrow \infty$ is generally easier to handle than a full relativistic treatment at finite speed of light c , which of course does not mean that Newtonian dynamics is more correct than relativity. Just as relativity compels us to think more carefully about how to formulate fundamental physical laws, the thermostistical analysis of finite systems forces us to pay more rigorous attention to mathematical and physical consistency criteria [7] in thermodynamics. This profound insight can be attributed to Gibbs, who wrote on page 179 of his fundamental treatise [3]: “It would seem that in general averages are the most important, and that they lend themselves better to analytical transformations. This consideration would give preference to the system of variables in which $\log V$ [= S_G in our notation] is the analogue of entropy. Moreover, if we make ϕ [= S_B in our notation] the analogue of entropy, we are embarrassed by the necessity of making numerous exceptions for systems of one or two degrees of freedoms.” Gibbs was, of course, well aware that statistical fluctuations become important in finite systems and that, therefore, the exact thermodynamic mean value relations (6) and (7) have to be complemented by detailed fluctuation analysis, as nowadays the norm in DNA and colloid experiments [21–23].

In this contribution, we have surveyed the notion of thermodynamic temperature in the microcanonical and the canonical ensemble. For isolated microcanonical systems, the Gibbs volume entropy fulfills exactly the standard laws of thermodynamics as well as equipartition for a wide range of systems, including all classical standard Hamiltonian systems regardless of their size. For finite systems, fluctuation analysis provides important physical information beyond the mean values that define standard thermodynamic state-variables. The microcanonical Gibbs formalism agrees with the Clausius relation [8], implies strictly non-negative temperatures and, hence, ensures Carnot efficiencies ≤ 1 . By contrast, the Boltzmann entropy, which can yield ‘negative absolute temperatures’, is not a consistent thermodynamic entropy if one adopts the standard Laws of Thermodynamics, as summarized in Eqs. (6)–(8). It is therefore not obvious to us why one should favor an entropy that can violate Planck’s law (8) over one that fulfills it rigorously.

Notwithstanding, the Boltzmann temperature plays an important role as an effective bath temperature in the canonical ensemble, describing a subsystem that is in weak contact with a quasi-infinite environment. If the bath behaves normally (e.g., ideal gas-like), then the Boltzmann temperature practically coincides with the Gibbs temperature. Subtle differences arise, however, for systems that are non-weakly coupled to an environment, as typically the case for nano-scale devices. In the presence of strong coupling, the specific heat of the device can become negative [28, 29] even though the total system consisting of device and bath is thermodynamically stable. This feature is in stark contrast to the weak-coupling case, where the canonical specific heat of the subsystem is strictly positive.

Last and not least, some authors [19, 20, 24] have recently criticized the microcanonical Gibbs formalism [3, 5] by limiting their discussion to infinite systems and advocating modified versions of the thermodynamic laws, tailored to favor their own preferred entropy definitions. If one accepts such reasoning, then one must also be willing to replace the exact Eqs. (6)–(8) with inexact approximations – which seems a steep price to pay. The exactness of Eqs. (6)–(8) is not a consequence of specific postulates but follows from basic integral and differential calculus (the ‘proofs’ are trivial and take only a few lines [7]). Hence, even if one dislikes the Gibbs formalism as developed in Refs. [3, 5, 6], one should at least acknowledge the correctness of the mathematically rigorous results (6)–(8). Moreover, instead of focusing on the discussion of abstract postulates [24], it may also be useful to remind ourselves that the purpose of any thermodynamic theory should be the prediction of physically measurable quantities, such as pressure, magnetization, etc. which correspond to operationally well-defined statistical averages. We would therefore encourage readers who care about the practical applicability of theoretical concepts to perform the following simple numerical experiment:

1. Place N_1 heavy particles (mass m_1) and N_2 light particles¹⁰ (mass $m_2 < m_1$) randomly on a finite 1D interval $[0, L]$, and assign to each particle some initial velocity, corresponding to some total energy $E = (1/2)(\sum_{i=1}^{N_1} m_1 v_i^2 + \sum_{j=1}^{N_2} m_2 v_j^2)$.
2. Evolve the system by assuming elastic momentum-conserving point-particle collisions and total reflection at the interval boundaries.

¹⁰ Two particle types are required to make the dynamics sufficiently ergodic, if $N = N_1 + N_2 > 1$.

3. Measure the kinetic temperatures $k_B \bar{T}_s = \langle m_s v_i^2 \rangle$ for each species $s = 1, 2$ and the kinetic pressure \bar{p} , conventionally defined as the mean momentum transfer to the interval boundaries per unit time, by taking standard time averages.
4. Compare \bar{T}_s and \bar{p} with predictions from the various entropy definitions for any combination (N_1, N_2) , starting with $(1, 0), (0, 1), (1, 1), \dots$

If one still intends to discard the Gibbs entropy afterwards, then one will have to explain why it is common sense to replace an entropy definition that produces correct predictions for all combinations (N_1, N_2) by another one that does not.

Acknowledgements The authors thank Michele Campisi, Gert-Ludwig Ingold and Peter Talkner for numerous insightful comments and discussions on these topics over the last ten years. This work was supported by the DFG Cluster of Excellence ‘Nanosystems Initiative Munich’ (P.H.) and DFG Cluster of Excellence ‘Origin and Structure of the Universe’ (S.H.) and an MIT Solomon Buchsbaum Fund Award (J.D.).

-
- [1] Kubo R. 1965. *Statistical Mechanics: An Advanced Course with Problems and Solutions*, Amsterdam NL, Elsevier B.V.. See Section 1.6 therein.
 - [2] Clausius R. 1865. Über verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie, *Ann. d. Phys.* **125**, 353-400.
 - [3] Gibbs JW. 1902. *Elementary Principles in Statistical Mechanics*, New York USA, Charles Scribner’s Sons.
 - [4] Planck M. 1906. *Vorlesungen über die Theorie der Wärmestrahlung*, Verlag von J. A. Barth, Leipzig, Germany, Chapter 4 (*Entropie und Wahrscheinlichkeit*), §136, p. 139.
 - [5] Planck M. 1957. *Theory of Heat; Vol. V, ‘Introduction to Theoretical Physics’*, The Macmillan Co., USA, 2-nd printing, translated by Brose H L. In particular, see §117 p. 226 and also Eq. (372) on p. 252, where the *Gibbs volume* $\Omega(E, Z)$ is selected to define the entropy.
 - [6] Hertz P. 1910. (a) Über die mechanischen Grundlagen der Thermodynamik. Part I. *Annalen der Physik* (33), 225-274; (b) Über die mechanischen Grundlagen der Thermodynamik. Part II. *Annalen der Physik* (33), 537-552.
 - [7] Hilbert S, Hänggi P, and Dunkel J. 2014. Thermodynamic Laws in Isolated Systems. *Phys. Rev. E* **90**, 062116.
 - [8] Campisi M. 2015. Construction of microcanonical entropy on thermodynamic pillars. *Phys. Rev. E* **91**, 052147.
 - [9] Schlüter A. 1948. Zur Statistik klassischer Gesamtheiten. *Z. Naturforsch.* **3A**, 350-360.
 - [10] Khinchin AI. 1949. *Mathematical Foundations of Statistical Mechanics*, New York, Dover Publ. Inc.
 - [11] Berdichevsky V, Kunin I and Hussain F. 1991. (a) Negative Temperature and Vortex Motion *Phys. Rev. A* **43**, 2050-2051. (b) *Phys. Rev. A* **44** 8439-8440.
 - [12] Campisi M. 2005. On the mechanical foundations of thermodynamics: The generalized Helmholtz theorem. *Studies in History and Philosophy of Modern Physics* **36**, 275-290.
 - [13] Dunkel J and Hilbert S. 2006. Phase transitions in small systems: microcanonical vs. canonical ensembles. *Physica A* **370**, 390-406.
 - [14] Casetti A, Kastner M and Nerattini R. 2009. Kinetic energy and microcanonical nonanalyticities in finite and infinite systems. *J. Stat. Mech.: Theory and Exp.* P07036.
 - [15] Junghans C, Bachmann M and Janke W. 2008. Thermodynamics of peptide aggregation processes: An analysis from perspectives of three statistical ensembles *J. Chem. Phys.* **128**, 085103.
 - [16] Dunkel J and Hilbert S. 2014. Consistent thermostatics forbids negative absolute temperatures. *Nat. Phys.* **10**, 67-72; note also the “Supplementary Information” therein.
 - [17] Baeten M and Naudts J. 2011. On the Thermodynamics of Classical Microcanonical Systems. *Entropy* **13**, 1186-1199.
 - [18] Sommerfeld A. 1951. *Vorlesungen über Theoretische Physik (Volume V): Thermodynamik und Statistik*, reprinted 2011, Verlag Harri Deutsch, Frankfurt am Main; see pp. 181-183 therein.
 - [19] Swendsen R H and Wang J S. 2014. Negative temperatures and the definition of entropy. arXiv:1410.4619.
 - [20] Wang J S. 2015. Critique of the Gibbs volume entropy and its implication. arXiv:1507.02022.
 - [21] Alemany A, Ribezzi, M, and Ritort F. 2013. Recent progress in fluctuation theorems and free energy recovery. In *Nonequilibrium Statistical Physics Today: Proceedings of the 11th Granada Seminar on Computational and Statistical Physics*. AIP Conference Proceedings **1332**, 96-110.
 - [22] A. Bérut *et al.* 2012. Experimental verification of Landauer’s principle linking information and thermodynamics. *Nature* **483**, 187-190.
 - [23] Roldan E and Martinez I A and Parrondo J M R and Petrov D. 2014. Universal features in the energetics of symmetry breaking. *Nat. Phys.* **10**: 457-461.
 - [24] Swendsen R H and Wang J S. 2015. The Gibbs” volume” entropy is incorrect. arXiv:1506.06911.
 - [25] Votyakov EV, Hidmi HI, De Martino A, and Gross D H E. 2002. Microcanonical Mean-Field Thermodynamics of Self-Gravitating and Rotating Systems. *Phys. Rev. Lett.* **89**, 031101.
 - [26] Callen H B. 1985. *Thermodynamics and an Introduction to Thermostatistics*. New York. John Wiley and Sons, revised edition. See Section 8 therein.

- [27] Campisi M, Talkner P, and Hänggi P. 2009. Finite bath fluctuation theorem. *Phys. Rev. E* **80**, 031145.
- [28] Hänggi P, Ingold G L and Talkner P. 2008. Finite quantum dissipation: the challenge of obtaining specific heat. *New J. Phys.* **10**, 115008.
- [29] Hänggi P, Ingold G L and Talkner P. 2009. Specific heat anomalies of open quantum systems. *Phys. Rev. E* **79**, 061105.
- [30] Campisi M, Talkner P and P Hänggi P. 2009. Thermodynamics and fluctuation theorems for a strongly coupled open quantum system: an exactly solvable case. *J. Phys. A: Math. Theor. (Fast Track)* **42**, 392002.
- [31] Rényi A. 1961. On measures of entropy and information. *Proc. Fourth Berkeley Symp. on Math. Statist. and Prob.* **1**: 547-561.
- [32] Baltes H P. 1976. *Spectra of Finite Systems: Review of Weyl's Problem*. Mannheim-Wien-Zürich, Bibliographisches Institut AG. ISBN 3-411-01491-1.