

Article

Thermodynamics, Statistical Mechanics and Entropy

Robert H. Swendsen

Physics Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA; swendsen@cmu.edu;
Tel.: +1-412-268-5211

Received: 30 September 2017; Accepted: 6 November 2017; Published: 10 November 2017

Abstract: The proper definition of thermodynamics and the thermodynamic entropy is discussed in the light of recent developments. The postulates for thermodynamics are examined critically, and some modifications are suggested to allow for the inclusion of long-range forces (within a system), inhomogeneous systems with non-extensive entropy, and systems that can have negative temperatures. Only the thermodynamics of finite systems are considered, with the condition that the system is large enough for the fluctuations to be smaller than the experimental resolution. The statistical basis for thermodynamics is discussed, along with four different forms of the (classical and quantum) entropy. The strengths and weaknesses of each are evaluated in relation to the requirements of thermodynamics. Effects of order $1/N$, where N is the number of particles, are included in the discussion because they have played a significant role in the literature, even if they are too small to have a measurable effect in an experiment. The discussion includes the role of discreteness, the non-zero width of the energy and particle number distributions, the extensivity of models with non-interacting particles, and the concavity of the entropy with respect to energy. The results demonstrate the validity of negative temperatures.

Keywords: thermodynamics; statistical mechanics; entropy

1. Introduction

Recently, the question of the proper definition of the thermodynamic entropy in statistical mechanics has been the subject of renewed interest. A controversy has arisen, which has revealed new insights into old issues, as well as unexpected disagreements in the basic assumptions of thermodynamics and statistical mechanics [1–27]. A significant point of contention is the validity of negative temperatures [28,29]. Although the controversy has been described by some as a choice between two microcanonical definitions of entropy, I believe that the important issues are much more fundamental, and that the microcanonical ensemble is at the root of much of the conflict.

A key area of disagreement is the structure of thermodynamics itself. The opposing groups have radically different views of thermodynamics, which must be reconciled before there is any hope of coming to an agreement on how the entropy should be defined within thermodynamics. If we can reach an agreement on thermodynamics and the criteria that the entropy should satisfy, we should be able to achieve a consensus on the proper definition of entropy.

The purpose of this paper is to present one view of those basic assumptions, while avoiding direct criticism of alternatives that have been suggested. However, because small effects of the order of $1/N$, where N is the number of particles, have played a significant role in References [1–11], my discussion will also take small effects seriously.

I present the structure of thermodynamics in Section 2 as a development of the views of Tisza [30] and Callen [31,32], as codified in Callen's well-known postulates of thermodynamics. I will argue that some of those postulates are unnecessarily restrictive, and I will abandon or modify them accordingly. Justifications for the remaining postulates will be presented, along with their consequences. Criteria that the thermodynamic entropy must satisfy arise directly from these postulates. It will

become clear that a definition of the entropy satisfying a minimal set of these criteria for the most general form of thermodynamics is not unique. However, additional criteria do make the entropy unique.

There is no need to limit thermodynamics to homogeneous systems or short-ranged interactions (within a system). On the other hand, it should be recognized that when thermodynamics is applied to the calculation of materials properties, it can be useful to assume homogeneity. Statistical definitions that give extensive entropy for homogeneous systems are then preferred, because the Euler equation is valid for such systems.

Thermodynamics is commonly taught by starting with the four Laws of Thermodynamics. In Section 3, I derive these laws from the postulates, to demonstrate the close connection between the laws and the postulates. The advantage of the postulates is that they make some necessary aspects of the structure of thermodynamics explicit.

For classical statistical mechanics, I start in Section 4 from the assumption that a physical system is in some definite microscopic state, but that we have only limited knowledge of what that state is. We can describe our knowledge of the microscopic state by a probability distribution, with the assumption that all states that are consistent with experimental measurements are equally likely. From this assumption, we can, in principle, calculate all thermodynamic properties without needing thermodynamics. A thermodynamic description of these properties must be consistent with calculations in statistical mechanics.

An advantage of defining the entropy in terms of probability, instead of either surfaces or volumes in phase space, is that Liouville's theorem does not prevent the entropy of an isolated system from increasing [33], and the limit of an infinite system is not necessary [34].

A striking difference between these two methods of calculating the properties of physical systems is that while statistical mechanics only predicts probability distributions, thermodynamics is deterministic. For this reason, I take the domain of thermodynamics to be limited to large, but finite systems. If a system has N particles, measurements of physical properties typically have relative uncertainties due to fluctuations of the order of $1/\sqrt{N}$. If we assume that the resolution of experimental measurements is not capable of measuring these fluctuations, then statistical mechanics provides a deterministic prediction. Since typical values of N are in the range of 10^{12} – 10^{24} , this criterion is sufficiently general to justify the wide application of thermodynamics to physical experiments. However, I will not take the ("thermodynamic") limit $N \rightarrow \infty$ (see Section 10).

Some workers in the field believe that thermodynamics should also apply to small systems—even a system consisting of a single particle [9,11]. Demanding that thermodynamics be valid for small systems would put an additional constraint on the structure of the theory, since thermodynamics must still be valid for large systems. Although thermodynamic ideas can be of value in understanding the properties of small systems, I believe that the full application of thermodynamics to small systems must be treated with care as a separate topic.

To find the proper definition of the entropy, it is natural to compare how thermodynamics and statistical mechanics make the same predictions of the results of macroscopic experiments. This consideration gives a candidate for the definition of entropy that satisfies the most basic criteria for consistent thermodynamics, the Boltzmann entropy [35–39]. It correctly predicts the mode of the probability distribution for equilibrium values of macroscopic experiments. Since the width of the distribution is $1/\sqrt{N}$ for most quantities, it is well within experimental resolution.

However, there are other criteria that the entropy might satisfy. These include exact adiabatic invariance and the prediction of the mean of an experimental observable rather than the mode, even though the difference is usually of order $1/N$. Indeed, some theorists have insisted that these criteria must be satisfied for thermodynamic consistency [3,4]. Although I do not think that they are necessary for consistency, I will include these additional criteria in the discussion to give a complete picture.

I will discuss four candidates for the entropy of classical systems in Sections 5–8.

For quantum systems, the discreteness of the spectrum of energy eigenvalues presents a special situation, which will be discussed in Section 9. A probability distribution over many states of the system, including linear combinations of eigenstates, is generated if the system of interest has ever been in thermal contact with another macroscopic system. The quantum microcanonical ensemble, which is only defined for the energy eigenvalues, is overly restrictive in that it omits all states that are formed as linear combinations of eigenstates. It has been shown that the properties of macroscopic quantum systems are continuous functions of the average energy, rather than being defined only at the energy eigenvalues [22].

The question of the form of the thermodynamic entropy in the description of a first-order phase transition is discussed in Section 10. An exact inequality requires that a plot of entropy vs. energy must be concave. However, opposition to concavity comes from the clear signature of a first-order transition found in the convexity of the Boltzmann entropy near a first-order transition [2,40–42]. This signature of a first-order transition is indeed very useful, but it is not present in the thermodynamic entropy.

Section 11 illustrates the differences between the various suggested forms for the entropy with systems that consist of non-interacting objects. In the limit that the components of a thermodynamic system are independent of each other, they should contribute independently to any extensive property. If all objects have the same properties, the entropy should be exactly proportional to their number. Almost all suggested forms for the entropy satisfy this criterion, but *not for finite systems*.

Negative temperatures are discussed in Section 12. They are found to be consistent with all thermodynamic requirements.

2. Thermodynamic Postulates

In this section, I deal with Callen's postulates for thermodynamics [31,32]. I first preview some conditions that Callen set forth for the postulates, then present the original set that he proposed. The postulates are then re-examined, first as reformulated in my textbook [43]. Finally, a minimal set of four essential postulates is proposed, along with three optional postulates, which are useful in certain situations.

2.1. Callen's Conditions

Early in his book [31,32], before presenting his postulates, Callen restricts attention to "simple systems," defined as

systems that are macroscopically homogeneous, isotropic, uncharged, and chemically inert, that are sufficiently large that surface effects can be neglected, and that are not acted on by electric, magnetic, or gravitational fields [31,32].

There are quite a few conditions in this sentence, not all of them necessary.

The condition that the systems "are sufficiently large that surface effects can be neglected," is both ambiguous and unnecessary. A gas in a container with adsorbing walls is subject to thermodynamic laws, but even a very large system would not be sufficient to neglect the walls, since there is always a condensation onto the walls at low temperatures. The systems can also be inhomogeneous and anisotropic, although care must be taken to generalize the concept of pressure.

If the system has a net charge, it will interact with other systems through the long-ranged electrical force, which should be excluded.

The condition that the systems be "chemically inert" is not necessary. Although I will present the thermodynamic equations for systems with just one chemical component, this is easily generalized. Chemical changes are described thermodynamically by the chemical potentials of the various components.

The condition that the systems "are not acted on by electric, magnetic, or gravitational fields," should not exclude static fields.

2.2. Callen's Postulates

I will give Callen's postulates along with some comments on their range of validity.

Callen's Postulate 1:

There exist particular states (called equilibrium states) of simple systems that, macroscopically, are characterized completely by the internal energy U , the volume V , and the mole numbers $N_1, N_2 \dots N_r$ of the chemical components [31,32].

For simplicity, I will only consider one type of particle. The generalization to r different types is straightforward, as is the generalization to include the magnetization, polarization, etc. I also use N to denote the number of particles, instead of the number of moles.

Callen's Postulate 2:

There exists a function (called the entropy S) of the extensive variables of any composite system, defined for all equilibrium states and having the following property: The values assumed by the extensive parameters in the absence of an internal constraint are those which maximize the entropy over the manifold of constrained equilibrium states [31,32].

This postulate is equivalent to the second law of thermodynamics in a very useful form. Since the entropy is maximized when a constraint is released, the total entropy cannot decrease, $\Delta S \geq 0$. The values of the released variables (energy, volume, or particle number) at the maximized entropy predict the equilibrium values.

This postulate also introduces the concept of a state function, which plays a significant role in the theory. In my reformulation, I have made the existence of state functions a separate postulate (see Sections 2.3 and 2.4).

Callen's Postulate 3:

The entropy of a composite system is additive over the constituent subsystems. The entropy is continuous and differentiable and is a monotonically increasing function of the energy [31,32].

The first sentence of this postulate, "The entropy of a composite system is additive over the constituent subsystems," is quite important. Together with Callen's Postulate 2, it requires two systems, j and k , in thermal equilibrium to satisfy the equation:

$$\left(\frac{\partial S_j}{\partial U_j} \right)_{V_j, N_j} = \left(\frac{\partial S_k}{\partial U_k} \right)_{V_k, N_k}. \quad (1)$$

The left side of Equation (1) depends only on the properties of system j , while the right side depends only on the properties of system k . The derivative $\partial S/\partial U$ clearly is a function of the temperature, and comparison with the classical ideal gas shows that function to be $\partial S/\partial U = 1/T$ [31,32,43].

The limitation to entropy functions that are "continuous and differentiable" is necessary for the use of calculus and justifiable, even for quantum systems with discrete spectra [22].

The limitation that the entropy is a "monotonically increasing function of the energy" is unnecessary. It is equivalent to the assumption of positive temperatures. It has the advantage of allowing the inversion of $S = S(U, V, N)$ to obtain $U = U(S, V, N)$. Callen uses Legendre transforms of $U = U(S, V, N)$ to express many of his results. However, the same thermodynamics can be expressed in terms of Legendre transforms of $S = S(U, V, N)$, without the limitation to positive temperatures.

Callen's Postulate 4:

The entropy of any system vanishes in the state for which $(\partial U/\partial S)_{V, N} = 0$ (that is, at the zero of temperature) [31,32].

This expression of the Third Law of Thermodynamics requires the entropy as a function of energy to be invertible, which was assumed in Callen's Postulate 3.

2.3. My Modifications of Callen's Postulates

In my 2012 textbook on statistical mechanics and thermodynamics, I reformulated Callen's postulates for mostly pedagogic reasons. They are given here for comparison, but I now think that they are somewhat too restrictive.

Modified Postulate 1: There exist equilibrium states of a macroscopic system that are characterized uniquely by a small number of extensive variables.

Modified Postulate 2: The values assumed by the extensive parameters of an isolated composite system in the absence of an internal constraint are those that maximize the entropy over the set of all constrained macroscopic states.

Modified Postulate 3: The entropy of a composite system is additive over the constituent subsystems.

Modified Postulate 4: The entropy is a monotonically increasing function of the energy for equilibrium values of the energy.

Modified Postulate 5: The entropy is a continuous and differentiable function of the extensive parameters.

Modified Postulate 6: The entropy is an extensive function of the extensive variables.

The Nernst Postulate was left out because it is not necessary for thermodynamics. It is only true for quantum systems, and is represented by the properties of quantum systems without including it as a postulate. Whether it is listed as a postulate is mainly a matter of taste.

2.4. The Essential Thermodynamic Postulates

As a result of recent developments, I have come to the conclusion that fewer postulates are required. The following are a list of the essential postulates that must be satisfied for consistent thermodynamics. I have included comments that explain why each is necessary.

Postulate 1: Equilibrium States

There exist equilibrium states of a macroscopic system that are characterized uniquely by a small number of extensive variables.

The term "state function" is used to denote any quantity that is a function of the small number of variables needed to specify an equilibrium state. The primary variables are the extensive ones, which specify quantities. Intensive variables, such as temperature, pressure, chemical potential, magnetic field, and electric field, are derived quantities as far as the postulates are concerned. Of course, after they are introduced into the thermodynamic formalism, they can also be used to characterize a thermodynamic equilibrium state through Legendre transforms.

Postulate 2: Entropy Maximization

The values assumed by the extensive parameters of an isolated composite system in the absence of an internal constraint are those that maximize the entropy over the set of all constrained macroscopic states.

This postulate is an explicit form of the second law of thermodynamics. It automatically specifies that the total entropy will not decrease if a constraint is released, and provides a way of calculating the equilibrium values of the new equilibrium state. In the second postulate, the importance of defining the entropy through the extensive variables becomes apparent.

Postulate 3: Additivity

The entropy of a composite system is additive over the constituent subsystems.

Additivity means that

$$S_{j,k}(E_j, V_j, N_j; E_k, V_k, N_k) = S_j(E_j, V_j, N_j) + S_k(E_k, V_k, N_k), \quad (2)$$

where S_j and S_k are the entropies of systems j and k . It should be noted that there is no extra condition that the joint entropy of two systems must satisfy. It also means that Equation (1) is generally valid, which implies the zeroth law of thermodynamics (see Section 3). Furthermore, it allows the identification of the inverse temperature as $\partial S/\partial E$, and allows the construction of a thermometer.

Postulate 3 means that the entropies of the individual systems, as functions of the extensive variables, are sufficient to perform any thermodynamic calculation.

Postulate 4, as given in my book [43] and above in the list of modified postulates (Section 2.3), is monotonicity. It has been demoted. Its purpose was to allow the inversion of the entropy as a function of energy to obtain energy as a function of entropy. Since this is not necessary, it will only be given below as a special case.

Postulate 4: Continuity and differentiability

The entropy is a continuous and differentiable function of the extensive parameters.

Some of the proposed forms of entropy do not satisfy this postulate, which has led to many workers to rely on the “thermodynamic limit,” in which the entropy per particle, S/N is computed as a function of U/N in the limit $N \rightarrow \infty$. I do not regard this limit as important, because the entropy, S , can be shown to be a continuous function of the energy, U , for finite systems[22]. The entropy of arbitrarily large finite systems can also differ qualitatively from the “thermodynamic limit” for first-order phase transitions [23,44].

This completes the minimal set of postulates that are necessary for consistent thermodynamics.

2.5. Optional Postulates

The remaining postulates are of interest for more restricted, but still important cases.

Postulate 5: Extensivity

The entropy is an extensive function of the extensive variables.

Extensivity means that

$$S(\lambda U, \lambda V, \lambda N) = \lambda S(U, V, N) \quad (3)$$

This postulate is not generally true, but it is useful for studying bulk material properties. If we are not interested in the effects of the boundaries, we can impose conditions that suppress them. For example, periodic boundary conditions will eliminate the surface effects entirely. There will still be finite-size effects, but they will usually be negligible away from a phase transition.

If the system is composed of N independent objects, the entropy is expected to be *exactly* extensive. This would be true of independent oscillators, N two-level objects, or a classical ideal gas of N particles. Unfortunately, most suggested expressions for the entropy in statistical mechanics do not give exactly extensive entropies for these systems, although the deviations from extensivity are small.

Extensivity implies that the Euler equation is valid.

$$U = TS - PV + \mu N \quad (4)$$

The Euler equation leads to the Gibbs–Duhem relation,

$$0 = S dT - V dP + N d\mu, \quad (5)$$

which shows that T , P , and μ are not independent for extensive systems.

Postulate 6: Monotonicity

The entropy is a monotonically increasing function of the energy for equilibrium values of the energy.

This postulate allows the inversion of $S = S(E, V, N)$ to obtain $U = U(S, V, N)$, and derive several thermodynamic potentials in a familiar form. The postulate is unnecessary, and Massieu functions can be used for a general treatment [22,23,44].

Postulate 7: Nernst Postulate

The entropy of any system is non-negative.

The Nernst postulate has not been phrased in terms of the limit $(\partial S/\partial U)_{V,N} \rightarrow \infty$, as is usual, since several theoretical models include spin glasses, which have non-zero (positive) entropy at zero temperature. This phrasing still implies that heat capacities and compressibilities vanish as $T \rightarrow 0$ (or, equivalently, as $\beta = 1/k_B T \rightarrow \infty$).

The Nernst Postulate is also known as the Third Law of Thermodynamics. It only applies to quantum systems, but, since all real systems are quantum mechanical, it ultimately applies to everything. It is listed as optional because classical systems are interesting in themselves and useful in understanding the special properties of quantum systems.

The Nernst postulate is often said to be equivalent to the statement that it is impossible to achieve absolute zero temperature. This is not correct, because reaching absolute zero is even more strongly prohibited by classical physics. An infinite amount of energy in addition to an infinite number of steps would be required to reach absolute zero if a system really obeyed classical mechanics [45]. The Nernst postulate only requires a finite amount of energy, but an infinite number of steps.

That completes the set of thermodynamic postulates, essential and otherwise. Before going on to discuss the explicit form of the entropy, I will discuss the derivation of the laws of the thermodynamics.

3. The Laws of Thermodynamics

The establishment of the laws of thermodynamics was extremely important to the development of the subject. However, they do not provide a complete basis for practical thermodynamics, which I believe the postulates do. Consequently, I prefer the postulates as providing an unambiguous prescription for the mathematical structure of thermodynamics and the required properties of the entropy.

The first law of thermodynamics is, of course, conservation of energy. It has been regarded as an inevitable assumption in virtually every field of physics, and, as a result, I have not stressed it, regarding it as obvious. It must be remembered that the identification of the phenomenon of heat as a manifestation of a form of energy was one of the great achievements of the nineteenth century. It is used throughout the thermodynamic formalism, most explicitly in the form,

$$dU = \delta Q + \delta W, \quad (6)$$

where dU is the differential change in the internal energy of a system, δQ is the heat added to the system, and δW is the work done on the system (the use of the symbol δ indicates an inexact differential).

The second law of thermodynamics follows directly from the second essential postulate. It is most concisely written as

$$\Delta S \geq 0, \quad (7)$$

where ΔS indicates a change of the total entropy upon release of any constraint.

The third law of thermodynamics is also known as the Nernst postulate. It has been discussed in Section 2.

The zeroth law of thermodynamics was the last to be named. It is found in a book by Fowler and Guggenheim [46] in the form

If two assemblies are each in thermal equilibrium with a third assembly, they are in thermal equilibrium with each other.

It was deemed to be of such importance that it was not appropriate to give it a larger number than the other laws, so it became the zeroth. It is a direct consequence of Equation (1), which follows from Postulates 2 and 3.

In Section 4, I begin the discussion of the relevant statistical mechanics with the calculation of the probability distribution as a function of the energies, volumes, and number of particles for a large number of systems that might, or might not, interact. I then discuss the Boltzmann and Gibbs entropies. Taking into account the width of the probability distribution of the energy and number of particles leads to the canonical entropy and the grand canonical entropy.

4. Macroscopic Probabilities for Classical Systems

The basic problem of thermodynamics is to predict the equilibrium values of the extensive variables after the release of a constraint between systems [31,43]. The solution to this problem in statistical mechanics does not require any assumptions about the proper definition of entropy.

Consider a collection of $M \geq 2$ macroscopic systems, which include all systems that might, or might not, exchange energy, volume, or particles. Denote the phase space for the j -th system by $\{p_j, q_j\}$, where (in three dimensions) p_j represents $3N_j$ momentum variables, and q_j represents $3N_j$ configuration variables. Making the usual assumption that interactions between systems are sufficiently short-ranged and may be neglected [22], the total Hamiltonian of the collection of systems can be written as a sum of contributions from each system.

$$H_T = \sum_{j=1}^M H_j(p_j, q_j) \quad (8)$$

The energy, volume, and particle number of system j are denoted as E_j , V_j , and N_j , and are subject to the conditions on the sums,

$$\sum_{j=1}^M E_j = E_T, \quad \sum_{j=1}^M V_j = V_T, \quad \sum_{j=1}^M N_j = N_T, \quad (9)$$

where E_T , V_T , and N_T are constants. I will only write the equations for a single type of particle. The generalization to a variety of particles is trivial, but requires indices that might obscure the essential argument. The systems do not overlap each other. Naturally, only $3(M-1)$ of the variables are independent.

I am not restricting the range of the interactions within any of the M systems. I am also not assuming homogeneity, so I do not, in general, expect extensivity [47]. For example, the systems might be enclosed by adsorbing walls.

Since I am concerned with macroscopic experiments, I assume that no measurements are made that might identify individual particles, whether or not they are formally indistinguishable [48]. Therefore, there are $N_T! / \left(\prod_{j=1}^M N_j!\right)$ different permutations for assigning particles to systems, and all permutations are taken to be equally probable.

The probability distribution for the macroscopic observables in equilibrium can then be written as

$$W(\{E_j, V_j, N_j\}) = \frac{1}{\Omega_T} \left(\frac{N_T!}{\prod_j N_j!} \right) \int dp \int dq \prod_{j=1}^M \delta(E_j - H_j). \quad (10)$$

The constraint that the N_j particles in system j are restricted to a volume V_j is implicit in Equation (10), and the walls containing the system may have any desired properties. Ω_T is a constant, which is determined by summing or integrating over all values of energy, volume, and particle number that are consistent with the values of E_T , V_T , and N_T in Equation (9). The value of the constant Ω_T does not affect the rest of the argument.

Equation (10) can also be written as

$$W(\{E_j, V_j, N_j\}) = \frac{1}{\Omega_T} \prod_{j=1}^M \Omega_j(E_j, V_j, N_j), \quad (11)$$

where

$$\Omega_j = \frac{1}{h^{3N_j} N_j!} \int_{-\infty}^{\infty} d^{3N_j} p_j \int d^{3N_j} q_j \delta(E_j - H_j), \quad (12)$$

and Ω_T is a normalization constant. The factor of $1/h^{3N_j}$, where h is Planck's constant, is included for agreement with the classical limit of quantum statistical mechanics [43].

5. The Boltzmann Entropy, S_B , for Classical Systems

Consider $M \geq 2$ systems with Hamiltonians $H_j(p_j, q_j)$. The systems are originally isolated, but individual constraints may be removed or imposed, allowing the possibility of exchanging energy, particles, or volume. The number M is intended to be quite large, since all systems that might interact are included. The magnitude of the energy involved in potential interactions between systems is regarded as negligible. The probability distribution for the extensive thermodynamic variables, energy (E_j), volume (V_j), and number of particles (N_j), is given by the expression in Equations (11) and (12) [49]. The logarithm of Equation (11) (plus an arbitrary constant, C) gives the Boltzmann entropy of the M systems.

$$S_T(\{E_j, V_j, N_j\}) = \sum_{j=1}^M S_{B,j}(E_j, V_j, N_j) - k_B \ln \Omega_T + C, \quad (13)$$

where the Boltzmann entropy for the j -th system is

$$S_{B,j}(E_j, V_j, N_j) = k_B \ln \Omega_j(E_j, V_j, N_j). \quad (14)$$

Using Equation (12),

$$S_{B,j} = k_B \ln \left[\frac{1}{h^{3N_j} N_j!} \int d^{3N_j} p \int d^{3N_j} q \delta(E_j - H_j(p, q)) \right]. \quad (15)$$

Since the total Boltzmann entropy is the logarithm of the probability $W(\{E_j, V_j, N_j\})$, maximizing the Boltzmann entropy is equivalent to finding the mode of the probability distribution. This is not the same as finding the mean of the probability distribution, but the difference between the mean and the mode is usually of order $1/N$. As mentioned in the Introduction, I will take even such small differences seriously.

5.1. Strengths of the Boltzmann Entropy

If any constraint(s) between any two systems is released, the probability distribution of the corresponding variable(s) is given by W in Equation (11). Since the Boltzmann entropy is proportional to the logarithm of W , it correctly predicts the mode of the probability distribution. Whenever the peak is narrow, the mode is a very good estimate for the mean since the relative difference is of the order

$1/N$. The mode is really used to estimate the measured value of some observable, which is within fluctuations of order $1/\sqrt{N}$.

5.2. Weaknesses of the Boltzmann Entropy

- $\Omega(E, V, N)$ has units of inverse energy, so that it is not proper to take its logarithm. This issue can easily be resolved by multiplying Ω by a constant with units of energy before taking the logarithm in Equation (15). Such a multiplicative factor adds an arbitrary constant to the individual entropies, but since it does not affect any thermodynamic prediction, it is not a serious problem.
- The Boltzmann entropy can be shown *not* to be adiabatically invariant [3,5,10]. Campisi provides the following definition: “A function $I(E, V)$ is named an adiabatic invariant if, in the limit of very slow variation of $V(t)$ namely as $t \rightarrow \infty$, $I(E(t), V(t)) \rightarrow \text{const.}$ ” The violation of adiabatic invariance for the Boltzmann entropy results in a relative error of the order of $1/N$. The error is very small, but it is a weakness in the theory. It is a consequence of the Boltzmann entropy predicting the mode instead of the average.
- The assumption of a microcanonical ensemble (that the probability distribution of the energy is a delta function) is, strictly speaking, incorrect. Any thermal contact with another system will leave the system of interest with an uncertainty in its energy. Since the width of the energy distribution is typically narrow, of order $1/\sqrt{N}$, where N is the number of particles, the approximation is usually regarded as reasonable. It will turn out that this approximation is responsible for some of the weaknesses in the theory and most of the disagreements in the literature.
- The number of particles N is discrete. The relative distance between individual points is very small, but nevertheless, the entropy is not a continuous, differentiable function of N , as assumed in thermodynamics. This discreteness is usually ignored.
- The assumption that the number of particles in our system is known *exactly* is never true for a macroscopic system. The width of the distribution of values of N is much larger than the separation of points, and the average value $\langle N \rangle$ would be better to use than N .
- The maximum of the Boltzmann entropy corresponds to the mode of the probability distribution—not the mean. This leads to small differences of order $1/N$. For example, a partial derivative of the Boltzmann entropy with respect to energy gives

$$U = (3N/2 - 1)k_B T, \quad (16)$$

instead of

$$U = (3N/2)k_B T \quad (17)$$

This error is unmeasurable for macroscopic systems [21,50], but it is a weakness of the theory. It is the basis of the argument against the validity of the Boltzmann entropy [3,4].

- Although the classical ideal gas is composed of N particles that do not interact with each other, the Boltzmann entropy is not exactly extensive (see Equation (47)). We are used to this lack of extensivity for finite N , but it is not quite correct.
- In the case of a first-order phase transition, the energy distribution is not narrow, and the microcanonical ensemble is not justified. At a first-order transition, a plot of the Boltzmann entropy against energy typically has a region of positive curvature, although a well-known thermodynamic requirement for stability states that the curvature must be negative [31,32,43].

$$\left(\frac{\partial^2 S}{\partial E^2} \right)_{V,N} < 0 \quad (18)$$

This is actually an advantage as a clear signal of a first-order transition [2,40–42]. It is, however, a serious flaw if the Boltzmann entropy is regarded as a candidate for the thermodynamic entropy. See Section 10.

6. The Gibbs Entropy, S_G , for Classical Systems

The Gibbs (or volume) entropy is defined by an integral over all energies less than the energy of the system [51–53]. It has the form

$$S_G = k_B \ln \left[\int_0^E \Omega(E', V, N) dE' \right] \quad (19)$$

6.1. Strengths of the Gibbs Entropy

- The integral in the definition of the Gibbs entropy in Equation (19) is dimensionless, so there is no problem in taking its logarithm.
- The Gibbs entropy can be shown to be adiabatically invariant [3,5,10].
- For the Gibbs entropy of classical systems with a monotonically increasing density of states, the predicted energy is exactly correct [4,5,7–11], although this is not true of quantum systems [22,23].

6.2. Weaknesses of the Gibbs Entropy

- The assumption of a microcanonical ensemble (that the probability distribution of the energy is a delta function) is incorrect.
- The assumption that the number of particles in a system is known *exactly* is incorrect.
- The number of particles N is discrete, and should be replaced by the continuous variable $\langle N \rangle$.
- Although the classical ideal gas is composed of N particles that do not interact with each other, the Gibbs entropy is not exactly proportional to N . This lack of extensivity of the Gibbs entropy is essentially the same as for the Boltzmann entropy.
- The Gibbs entropy also violates the thermodynamic inequality,

$$\left(\frac{\partial^2 S}{\partial E^2} \right)_{V,N} < 0, \quad (20)$$

at a first-order transition (see Section 10).

- For a non-monotonic density of states, the Gibbs entropy gives counter-intuitive results. Consider two homogeneous systems with the same composition and a decreasing density of states for the energies of interest (for example, independent spins in a field, neglecting the kinetic energy). Let them have the same energy per particle, but let one be twice as large as the other. There will be no net transfer of energy if the two systems are put in thermal contact. The Boltzmann temperature will be the same for both systems, as expected. However, the Gibbs temperature of the larger system will be higher.
- For $\partial\Omega/\partial E < 0$, equal Gibbs temperatures do not predict that there will be no net energy flow if the two systems are brought into thermal contact.
- Because larger systems have higher Gibbs temperature, it is impossible to construct a thermometer that measures the Gibbs temperature in an energy range with a decreasing density of states.

7. The Canonical Entropy, S_C , for Classical Systems

This section presents a definition of entropy that makes essential use of the canonical ensemble. It will first be presented in Section 7.1 under the usual assumption that the system of interest is in thermal equilibrium, or has been in thermal equilibrium, with a much larger system. The justification

for applying the same equation for the entropy of a system that has been in thermal contact with a smaller system is given in Section 7.2.

7.1. The General Derivation of the Canonical Entropy

I have chosen to express the thermodynamic results in terms of Massieu functions [31], because they do not require the inversion of the fundamental relation $S = S(U, V, N)$ to find $U = U(S, V, N)$. Such an inversion is unnecessary and is not valid for systems with a non-monotonic density of states. The following uses the same notation as Callen for the Legendre transforms. The more familiar transform of $U = U(S, V, N)$ with respect to temperature is denoted by $U[T] = U - TS = F$, which is a function of T , V , and N .

Define a dimensionless entropy as

$$\tilde{S} = \frac{S}{k_B} \quad (21)$$

Since

$$dU = TdS - PdV + \mu dN \quad (22)$$

and the inverse temperature is $\beta = 1/k_B T$, we also have

$$d\tilde{S} = \beta dU + (\beta P)dV - (\beta\mu)dN, \quad (23)$$

where P is the pressure, V is the volume, μ is the chemical potential, and N is the number of particles. From Equation (23),

$$\beta = \left(\frac{\partial \tilde{S}}{\partial U} \right)_{V,N}. \quad (24)$$

The Legendre transform (Massieu function) of \tilde{S} with respect to β is given by

$$\tilde{S}[\beta] = \tilde{S} - \beta U = -\beta(U - TS) = -\beta F, \quad (25)$$

so that

$$\tilde{S}[\beta] = \ln Z(\beta, V, N). \quad (26)$$

The differential of the Massieu function $\tilde{S}[\beta]$ is

$$d\tilde{S}[\beta] = -Ud\beta + (\beta P)dV - (\beta\mu)dN. \quad (27)$$

This immediately gives

$$\left(\frac{\partial \tilde{S}[\beta]}{\partial \beta} \right)_{V,N} = -U. \quad (28)$$

To obtain \tilde{S} from $\tilde{S}[\beta]$, use

$$\tilde{S} = \tilde{S}[\beta] + \beta U, \quad (29)$$

and substitute $\beta = \beta(U)$.

7.2. The Justification of the Canonical Entropy

The use of the canonical ensemble to calculate the entropy of a general system requires some discussion. That the probability distribution of the energy is not a delta function has been proven [22]. If the system of interest has been in thermal contact with a much larger system, it is clear that the canonical ensemble is appropriate. However, if the system has instead been in thermal contact with a system that is the same size or even smaller, it is known that the distribution is narrower than the canonical distribution [54]. Nevertheless, the canonical entropy is appropriate for calculating the thermodynamic entropy.

Consider three macroscopic systems labeled A , B , and C . Let systems A and B be constructed to be the same, and in particular to be equal in size. Let system C be much larger than A and B .

Suppose all three systems are in thermal contact and have come to equilibrium. The entropies of systems A and B are then equal, and given by the canonical form discussed in the previous subsection.

Now remove system C from thermal contact with the other two systems, which is obviously a reversible process. A or B are still in equilibrium with each other at the same temperature as before C was removed. For consistency, the entropies must be unchanged. If the entropy were to decrease, it would be a violation of the second law of thermodynamics (and the second essential postulate). If the entropy were to increase upon separation, putting system C back into thermal contact with systems A and B would decrease the entropy, which would also violate the second law. The only possibility consistent with the second law is that the entropy is unchanged. Therefore, the canonical entropy is properly defined for all systems and sizes, regardless of their history.

7.3. Strengths of the Canonical Entropy

The canonical entropy for the thermodynamic entropy is superior to either the Boltzmann or the Gibbs entropies.

- The canonical entropy is adiabatically invariant.
- The canonical entropy gives a thermodynamically correct description of first-order transitions (see Section 10).
- The canonical entropy has an energy term of $(3/2)k_B N \ln(U/N)$ for the ideal gas, which is correct. As a consequence, the canonical entropy also gives the exact energy for the classical ideal gas.

$$U = (3N/2)k_B T \quad (30)$$

7.4. Weaknesses of the Canonical Entropy

- The number of particles N is discrete.
- The canonical entropy assumes that the value of N is known exactly, instead of using a distribution of possible values of N .
- The deviation from exact extensivity (in the factor $1/N!$) is a weakness. This lack of extensivity of the canonical entropy differs from that of the Boltzmann entropy, in that it does not affect the energy-dependent term.

8. The Grand Canonical Entropy, S_{GC} , for Classical Systems

The grand canonical entropy satisfies all criteria required by thermodynamics. It is calculated in much the same way as the canonical entropy.

8.1. The Definition of the Grand Canonical Legendre Transform

For the grand canonical ensemble, $S[\beta, (\beta\mu)]$ will be used. I have put parentheses around the second variable to emphasize that the product of β and μ is to be treated as a single variable. To find the Legendre transform with respect to both β and $(\beta\mu)$ use the equation

$$-(\beta\mu) = \left(\frac{\partial \tilde{S}}{\partial N} \right)_{U, N} . \quad (31)$$

in addition to Equation (24).

The Legendre transform (Massieu function) of \tilde{S} with respect to both β and $(\beta\mu)$ is given by

$$\tilde{S}[\beta, (\beta\mu)] = \tilde{S} - \beta U + (\beta\mu)N, \quad (32)$$

so that

$$\tilde{S}[\beta, (\beta\mu)] = \ln \mathcal{Z}(\beta, V, (\beta\mu)), \quad (33)$$

where \mathcal{Z} is the grand canonical partition function.

The differential of the Massieu function $\tilde{S}[\beta, (\beta\mu)]$ is

$$d\tilde{S}[\beta, (\beta\mu)] = -Ud\beta + (\beta P)dV + Nd(\beta\mu). \quad (34)$$

This immediately gives

$$\left(\frac{\partial \tilde{S}[\beta, (\beta\mu)]}{\partial \beta} \right)_{V, (\beta\mu)} = -U, \quad (35)$$

and

$$\left(\frac{\partial \tilde{S}[\beta, (\beta\mu)]}{\partial (\beta\mu)} \right)_{\beta, V} = N. \quad (36)$$

To obtain \tilde{S} from $\tilde{S}[\beta, (\beta\mu)]$, use

$$\tilde{S} = \tilde{S}[\beta, (\beta\mu)] + \beta U - (\beta\mu)N, \quad (37)$$

and replace the β and $(\beta\mu)$ dependence by U and $\langle N \rangle$.

8.2. Strengths of the Grand Canonical Entropy

The grand canonical entropy retains the advantages of the canonical entropy, but also has a correct description of the distribution of particles. It provides a completely consistent description of the properties of a thermodynamic system.

The grand canonical entropy for the classical ideal gas is exactly extensive, which is expected of a model in which there are no explicit interactions between particles.

8.3. Weaknesses of the Grand Canonical Entropy

None.

9. Quantum Statistical Mechanics

Most of the considerations discussed above in the previous five sections on the entropy of classical systems are still applicable for the entropy of quantum systems. However, some features are new.

The Nernst postulate (optional postulate 7, see Section 2.5) is, of course, now applicable. This has the immediate effect that all expressions for the entropy are non-negative, and heat capacities and compressibilities vanish as $T \rightarrow 0$.

9.1. Quantum Boltzmann and Gibbs Entropies

Quantum mechanics brings a new aspect to the question of the correct definition of entropy. The energy spectrum of a finite quantum system is discrete, even though the density of states is a continuous function of the average energy when all linear combinations of the eigenstates are included [22]. The usual procedure is to define the Boltzmann entropy through the spectrum of eigenstates [38,39]. The fourth essential postulate, which requires continuity and differentiability (as does Callen's third postulate), then eliminates it as a candidate for the thermodynamic entropy. The discreteness of the energy eigenvalue spectrum also eliminates the quantum Gibbs entropy.

9.2. Pathological Hamiltonians

The Boltzmann entropy also fails if the Hamiltonian is sufficiently complicated that every energy level is non-degenerate [55]. In that case, S_B would vanish at every energy level because $\ln 1 = 0$. Even if the Hamiltonian were not so extremely pathological that every eigenstate was non-degenerate,

sufficient complexity could distort S_B . Such Hamiltonians have not (yet) played an explicit role in statistical mechanics, but the possibility of their existence is an argument against the quantum Boltzmann entropy.

The Gibbs entropy is not affected by such pathological Hamiltonians, nor are the canonical or grand canonical entropies.

9.3. Quantum Grand Canonical Ensemble

The grand canonical ensemble includes all microscopic states, as well as the width of the energy and particle number distributions. When computing the grand canonical (or canonical) partition function, it is sufficient to sum over all energy eigenvalues, since cross terms arising from linear combinations of eigenstates vanish in equilibrium [43]. The derivation is found in most textbooks, and is essentially the same as given in Section 8 for classical systems. Applications of the quantum grand canonical ensemble include the standard treatments of Fermi–Dirac and Bose–Einstein gases.

9.4. Quantum Canonical Ensemble

For quantum systems that do not vary the particle number, such as spin systems, the canonical ensemble is more appropriate. The derivation of the canonical ensemble is the same as for the classical case. Details are given in Ref. [22], which also contains explicit derivations for the entropy of a system of independent quantum simple harmonic oscillators and a system of two-level objects. In addition to these expressions for the entropy being continuous functions of the energy, they are also exactly extensive, not just in the limit of $N \rightarrow \infty$.

10. First-Order Phase Transitions

Systems that exhibit first-order transitions require careful treatment. Strictly speaking, if a phase transition is defined as a non-analyticity of the thermodynamic potentials, finite systems do not have phase transitions. However, if a system is large enough, it can change its energy by a significant amount over an extremely small temperature range, which is indistinguishable experimentally from a true first-order transition.

The density of states (or the Boltzmann entropy) is a valuable tool to identify such quasi-first-order behavior [40–42]. It shows the presence of a quasi-first-order transition by a region of convexity as a function of the energy. However, it will be proven in the next subsection that the entropy must be concave, making the Boltzmann (or the Gibbs) entropy unsuitable candidates for the thermodynamic entropy.

10.1. Why the Entropy Must Be Concave

To derive the well-known inequality in Equation (18) from the second law of thermodynamics [31,32,43], consider two systems, A and B , that are isolated and placed in thermal equilibrium with each other. The equilibrium energies are U_A and U_B . By the second law, any shift of energy, ΔU , must not decrease the total entropy. Suppressing the V - and N -dependence,

$$S_A(U_A) + S_B(U_B) \geq S_A(U_A + \Delta U) + S_B(U_B - \Delta U). \quad (38)$$

The volume and the particle number are held constant in Equation (38).

Now let the two systems be constructed identically, so that $U_A = U_B = U$ and $S_A(U) = S_B(U) = S(U)$. Then, Equation (38) becomes

$$2S(U) \geq S(U + \Delta U) + S(U - \Delta U) \quad (39)$$

This equation is valid for all ΔU , large or small. If we let $\Delta U \rightarrow 0$,

$$0 \geq \lim_{\Delta U \rightarrow 0} \left[\frac{S(U + \Delta U) + S(U - \Delta U) - 2S(U)}{(\Delta U)^2} \right] = \frac{\partial^2 S}{\partial U^2}. \tag{40}$$

Making the constraint of constant volume and particle number explicit, this gives the inequality

$$\left(\frac{\partial^2 S}{\partial U^2} \right)_{V,N} \leq 0. \tag{41}$$

10.2. Representation of First-Order Transitions by the Canonical Entropy

To see the effect of using the canonical entropy, a simple generic model density of states can be constructed that has this property.

$$\ln \Omega_1(E) = AN \left(\frac{E}{N} \right)^\alpha + BN \exp \left(-\frac{E_{N,0}^2}{2\sigma_N^2} \right) - BN \exp \left(-\frac{(E - E_{N,0})^2}{2\sigma_N^2} \right) \tag{42}$$

The center of the Gaussian term is taken to be $E_{N,0} = fN\epsilon$, and the width of the Gaussian is $\sigma_N = gE_{N,0}$. If the parameter B does not depend on N , this model corresponds to a mean-field transition, while if B decreases as N increases, it corresponds to a system with short-range interactions. The behavior is qualitatively the same in both cases. I will treat the case of B being constant, because it is the more stringent test of the method (see also Section 10.3).

Figure 1 shows $S_B(E)/N = k_B \ln \Omega(E)/N$ vs. $E/N\epsilon$ as a solid curve. The dip in the density of states and the region of positive curvature can be seen clearly. Figure 1 also shows $S_C(U)/N$ for $N = 10, 50$, and 250 , where $U = \langle E \rangle$. In all cases, the plot of S_C shows the negative curvature required for stability. For quantum models that show the same effects, see Ref. [23].

It is apparent from Figure 1 why S_B is favored in a numerical analysis of first-order phase transitions, even though it is not the thermodynamic entropy. S_C vs. U provides no clear signal for the presence of the transition, just a very nearly straight section of the plot.

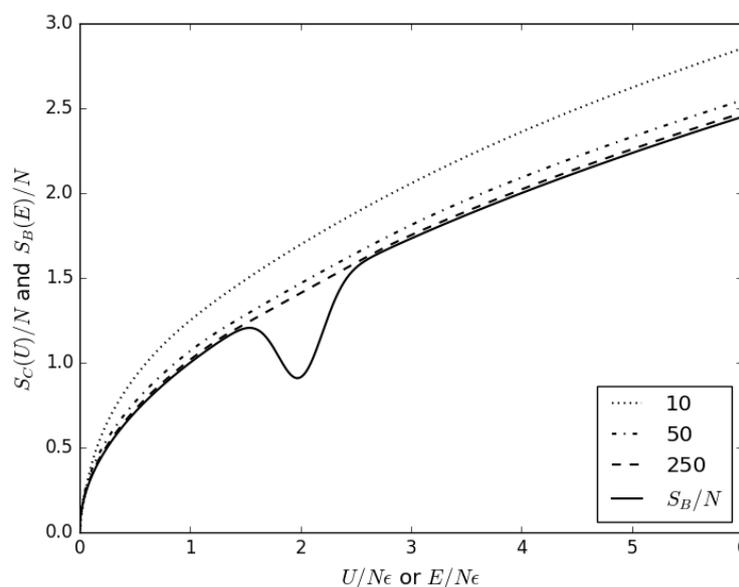


Figure 1. The top curves (dotted, dot-dash, and dashed) indicate the canonical entropy S_C/N for $N = 10, 50$, and 250 . The solid curve shows S_B , which is the same in the model for all N . The parameters are $\epsilon = 1, \alpha = 0.5, A = 1, B = 0.4, f = 2$, and $g = 0.1$, but other values give similar results.

10.3. The “Thermodynamic” Limit

The “thermodynamic” limit takes various thermodynamic functions (in particular, the entropy), divides them by the total number of particles (or something proportional to it), and takes the limit of $N \rightarrow \infty$ [56–60]. I believe that this procedure can give a distorted picture of the behavior of thermodynamic systems.

First of all, the application of thermodynamics to a real experiment requires the full entropy, S , not the entropy per particle, S/N . When two systems at different temperatures are put in thermal contact, the final temperature will depend on the relative sizes.

Next, if there is a qualitative difference between the behavior of the finite system and the thermodynamic limit, the finite system takes precedence. There is such a qualitative difference in the Boltzmann entropy for systems that show a first-order transition. The region of positive curvature in S_B diverges as $N^{(d-1)d} = N^{1-1/d}$ for short-range forces. Dividing by N gives $N^{-1/d}$, which goes to zero as $N \rightarrow \infty$. Therefore, in the limit, there is no positive curvature in S_B/N . However, even for a very large system, the Boltzmann entropy has a section of positive curvature [40–42].

There is an additional problem if the system has long-range interactions. Then the region of positive curvature in S_B goes as N , rather than $N^{1-1/d}$. The region of positive curvature does not become a straight line in this case, and even violates the thermodynamic inequality in Equation (41) in the limit $N \rightarrow \infty$ [58]. To exclude mean field models on this basis is unnecessary. As the example in Figure 1 shows, S_C is always concave, and there is no need to treat mean field models as exceptions.

11. Collections of Non-Interacting Objects

In the limit that the elementary objects that constitute a system do not interact with each other, all properties are expected to be exactly proportional to the number of objects. Examples include non-interacting classical particles, non-interacting spins, and independent simple harmonic oscillators. The entropy of such a system is expected to be extensive. In fact, extensivity is a consequence of the third essential postulate (Additivity).

11.1. The Classical Ideal Gas

To provide a simple comparison between the four expressions for the entropy, I will give the predictions of each for the entropy of the classical ideal gas. Since the ideal gas is defined in the limit of no interactions between the particles, every particle is expected to make exactly the same contribution to the entropy. For fixed E/N and V/N , there should be the same S/N . That is, the entropy should be exactly extensive.

To calculate the four forms of entropy, the density of states of the classical ideal gas can be evaluated as

$$\Omega_{CIG}(E, V, N) = \frac{\pi^{3N/2}}{(3N/2 - 1)!} \frac{1}{h^{3N} N!} V^N (2m)^{3N/2} E^{3N/2-1}. \tag{43}$$

11.1.1. The Boltzmann Entropy, S_B

S_B is simply related to the logarithm of Equation (43).

$$S_B(E, V, N) = k_B \ln \left(\frac{\pi^{3N/2}}{(3N/2 - 1)!} \frac{1}{h^{3N} N!} V^N (2m)^{3N/2} E^{3N/2-1} \right), \tag{44}$$

or

$$S_B(E, V, N) = k_B \left[\ln \left(\frac{E^{3N/2-1}}{(3N/2 - 1)!} \right) + \ln \left(\frac{V^N}{N!} \right) + N \ln X \right], \tag{45}$$

where X is a constant.

$$X = \frac{\pi^{3/2}}{h^3} (2m)^{3/2} \tag{46}$$

The Boltzmann entropy has terms that contain $(3N/2 - 1)!$ and $N!$, which prevent it from being exactly extensive. The exponent of E is $3N/2 - 1$, which also violates exact extensivity.

11.1.2. The Gibbs Entropy, S_G

To calculate the Gibbs entropy for the ideal gas, we must first integrate the density of states. The result for the classical ideal gas is

$$S_G(E, V, N) = k_B \left[\ln \left(\frac{E^{3N/2}}{(3N/2)!} \right) + \ln \left(\frac{V^N}{N!} \right) + \ln X^N \right], \quad (47)$$

where X is the same constant.

The exponent is now equal $3N/2$, which gives the correct energy. The factors of $1/(3N/2)!$ and $1/N!$ are still not exactly extensive.

11.1.3. The Canonical Entropy, S_C

Finding the canonical entropy involves a bit more algebra, but it is still straightforward. The result is

$$S_C = k_B \left[N \left(\frac{3}{2} \right) \ln \left(\frac{U}{N} \right) + \ln \left(\frac{V^N}{N!} \right) + N \left(\frac{3}{2} \right) \ln \left(\frac{4\pi m}{3h^2} \right) + \frac{3N}{2} \right]. \quad (48)$$

Note that Equation (48) uses $U = \langle E \rangle$. The factorials associated with the energy integral over the (surface or volume of the) sphere in momentum space, appear in the Boltzmann and Gibbs entropies, but not in the canonical entropy. Only the term associated with the volume involves a factorial ($1/N!$).

11.1.4. The Grand Canonical Entropy of the Ideal Gas

The calculation of the grand canonical entropy of the classical ideal gas proceeds in much the same way as that of the canonical. The result is

$$S_{GC} = \langle N \rangle k_B \left[\frac{3}{2} \ln \left(\frac{U}{\langle N \rangle} \right) + \ln \left(\frac{V}{\langle N \rangle} \right) + \ln \left(\frac{4\pi m}{3h^2} \right)^{3/2} + \frac{5}{2} \right] \quad (49)$$

This expression for the entropy of a classical ideal gas is exactly extensive [47], and no use has been made of Stirling's approximation.

11.2. Quantum Independent Simple Harmonic Oscillators

A system of N simple harmonic oscillators, each with frequency ω , has been treated in Ref. [22]. The energy levels of the k -th oscillator are

$$E_k(n_k) = \hbar\omega(n_k + 1/2), \quad (50)$$

where \hbar is Planck's constant and $n_k = 0, 1, 2, \dots$

The degeneracy of the energy level with a total energy of $E_n = \hbar\omega[n + (1/2)N]$ is $(n + N - 1)!/n!(N - 1)!$, and the Boltzmann entropy is

$$S_{B,SHO} = k_B \ln \Omega = k_B \ln \left[\frac{(n + N - 1)!}{n!(N - 1)!} \right]. \quad (51)$$

This expression is clearly a discrete function of the energy.

For the canonical entropy, define a dimensionless energy variable

$$x = \hat{U}_{\text{SHO}}/N\hbar\omega = \frac{U_{\text{SHO}} - \frac{1}{2}\hbar\omega N}{N\hbar\omega}, \quad (52)$$

and write the canonical entropy as

$$S_{\text{C,SHO}} = Nk_B [-x \ln x + (1+x) \ln(1+x)]. \quad (53)$$

This expression for the canonical entropy is clearly extensive and a continuous function of the energy.

$S_{\text{B,SHO}}$, $S_{\text{G,SHO}}$, and $S_{\text{C,SHO}}$ differ for all finite systems. However, the expressions for the entropies per SHO, $S_{\text{B,SHO}}/N$, $S_{\text{G,SHO}}/N$, and $S_{\text{C,SHO}}/N$, agree in the limit $N \rightarrow \infty$.

11.3. Quantum, Non-Interacting, Two-Level Objects

The total energy for the eigenstates of N two-level quantum objects is,

$$H_{2\text{-level}} = \epsilon \sum_{k=1}^N n_k, \quad (54)$$

where ϵ is the energy difference between the two levels in each object, and $n_k = 0$ or 1 . If $n = \sum_{k=1}^N n_k$, the degeneracy of the n -th energy level is given by $N!/ [n!(N-n)!]$, so that the Boltzmann entropy is

$$S_{\text{B,2-level}} = k_B \ln \left(\frac{N!}{n!(N-n)!} \right). \quad (55)$$

The canonical entropy of a system of non-interacting two-level objects has been found in Ref. [22]. Define a dimensionless energy for the spin system by

$$y = \frac{U_{2\text{-level}}}{N\epsilon}, \quad (56)$$

where $0 \leq y \leq 1$. The canonical entropy is then given by

$$S_{\text{C,2-level}} = -Nk_B [y \ln y + (1-y) \ln(1-y)]. \quad (57)$$

$S_{\text{C,2-level}}$ has positive temperatures for $y < 1/2$, and negative temperatures for $y > 1/2$, as expected, with an obvious symmetry for $y \leftrightarrow 1-y$. The entropy goes to zero in the limits $y \rightarrow 0$ and $y \rightarrow 1$, also as expected.

The Boltzmann entropy per two-level object, $S_{\text{B,2-level}}/N$, agrees with the canonical entropy, $S_{\text{C,2-level}}/N$, in the limit of $N \rightarrow \infty$. For any finite system, they are different. In particular, the N -dependence of $S_{\text{B,2-level}}/N$ is discrete.

The Gibbs entropy per two-level object agrees with the canonical entropy for $y \leq 1/2$ in the limit of $N \rightarrow \infty$, but is completely different for $y > 1/2$, where $\lim_{N \rightarrow \infty} S_{\text{G,2-level}}/N = k_B \ln 2$.

12. Negative Temperatures

Several quantum models display negative temperatures [28,29]: The two-level objects discussed in Section 11.3, Ising models, Potts models, xy- and rotator models, Heisenberg models, etc. They all exhibit a density of states which decreases for high energies. The natural way to describe such models is in terms of the inverse temperature, $\beta = k_B T$, which goes from $+\infty$ ($T = 0$), through zero ($T = \infty$), to $-\infty$ ($T = -0$). If these models are represented by the canonical or grand canonical entropy, they are consistent with all thermodynamic principles.

13. Discussion

In the course of this paper, I have presented a view of thermodynamics that is close to that found in most science and engineering textbooks. The differences lie mainly in the relaxation of the demand for extensivity, and dropping the requirement of monotonicity of the entropy as a function of the energy. The first is to allow the treatment of inhomogeneous systems, such as a gas contained in adsorbing walls, and the second is an unnecessary restriction to positive temperatures. The definition of the entropy is also expressed through the probability distributions of the extensive variables, which avoids having Liouville's theorem prevent the entropy of an isolated system from increasing [33]. This view of thermodynamics is not universally accepted [9,11], but I have tried to address issues of which I am aware.

For completeness, I have followed the advocates of opposing views [3–11] in considering very small ($1/N$) effects in comparing different forms of the entropy. Nevertheless, I do not agree that such small effects are of significance. If $1/N$ -effects are ignored, along with the discreteness of particle number and quantum energy, the Boltzmann entropy satisfies all criteria except for pathological quantum models (see Section 9) and the requirement of concavity at first-order transitions.

The Gibbs entropy does not have a difficulty with pathological models; it does have the same problem with concavity as the Boltzmann entropy, and it fails to give the correct equilibrium for a density of states that decreases with energy [19,21]. Ultimately, the last issue is the most critical in eliminating the Gibbs entropy as a candidate for the thermodynamic entropy. Naturally, this conclusion rests on the acceptance of the postulates of thermodynamics in Section 2.4, which is still a subject of debate.

The grand canonical entropy satisfies all criteria, answering the objections that led to questioning the thermodynamic consistency of negative temperatures. The validity of the concept of negative temperatures is confirmed.

Conflicts of Interest: The author declares no conflict of interest.

References and Notes

- Berdichevsky, V.; Kunin, I.; Hussain, F. Negative temperature of vortex motion. *Phys. Rev. A* **1991**, *43*, 2050–2051.
- Gross, D.H.E.; Kenney, J.F. The microcanonical thermodynamics of finite systems: The microscopic origin of condensation and phase separations, and the conditions for heat flow from lower to higher temperatures. *J. Chem. Phys.* **2005**, *122*, 224111.
- Campisi, M. On the mechanical foundations of thermodynamics: The generalized Helmholtz theorem. *Stud. Hist. Philos. Mod. Phys.* **2006**, *36*, 275–290.
- Dunkel, J.; Hilbert, S. Phase transitions in small systems: Microcanonical vs. canonical ensembles. *Physica A* **2006**, *370*, 390–406.
- Campisi, M.; Kobe, D.H. Derivation of the Boltzmann principle. *Am. J. Phys.* **2010**, *78*, 608–615.
- Romero-Rochin, V. Nonexistence of equilibrium states at absolute negative temperatures. *Phys. Rev. E* **2013**, *88*, 022144.
- Dunkel, J.; Hilbert, S. Consistent thermostats forbids negative absolute temperatures. *Nat. Phys.* **2014**, *10*, 67–72.
- Dunkel, J.; Hilbert, S. Reply to Frenkel and Warren. *arXiv* **2014**, arXiv:1403.6058v1.
- Hilbert, S.; Hänggi, P.; Dunkel, J. Thermodynamic laws in isolated systems. *Phys. Rev. E* **2014**, *90*, 062116.
- Campisi, M. Construction of microcanonical entropy on thermodynamic pillars. *Phys. Rev. E* **2015**, *91*, 052147.
- Hänggi, P.; Hilbert, S.; Dunkel, J. Meaning of temperature in different thermostats ensembles. *Philos. Trans. R. Soc. A* **2016**, *374*, 20150039.
- Miranda, E.N. Boltzmann or Gibbs Entropy? Thermostatistics of Two Models with Few Particles. *Int. J. Mod. Phys.* **2015**, *6*, 1051–1057.
- Frenkel, D.; Warren, P.B. Gibbs, Boltzmann, and negative temperatures. *Am. J. Phys.* **2015**, *83*, 163–170.
- Vilar, J.M.G.; Rubi, J.M. Communication: System-size scaling of Boltzmann and alternate Gibbs entropies. *J. Chem. Phys.* **2014**, *140*, 201101.

15. Schneider, U.; Mandt, S.; Rapp, A.; Braun, S.; Weimer, H.; Bloch, I.; Rosch, A. Comment on ‘Consistent thermostatics forbids negative absolute temperatures’. *arXiv* **2014**, arXiv:1407.4127v1.
16. Anghel, D.V. The stumbling block of the Gibbs entropy: The reality of the negative absolute temperatures. *arXiv* **2015**, arXiv:1509.08020v1.
17. Cerino, L.; Puglisi, A.; Vulpiani, A. Consistent description of fluctuations requires negative temperatures. *arXiv* **2015**, arXiv:1509.07369v1.
18. Poulter, J. In defense of negative temperature. *Phys. Rev. E* **2015**, *93*, 032149.
19. Swendsen, R.H.; Wang, J.S. The Gibbs volume entropy is incorrect. *Phys. Rev. E* **2015**, *92*, 020103(R).
20. Wang, J.S. Critique of the Gibbs volume entropy and its implication. *arXiv* **2015**, arXiv:1507.02022.
21. Swendsen, R.H.; Wang, J.S. Negative temperatures and the definition of entropy. *Physica A* **2016**, *453*, 24–34.
22. Swendsen, R.H. Continuity of the entropy of macroscopic quantum systems. *Phys. Rev. E* **2015**, *92*, 052110.
23. Matty, M.; Lancaster, L.; Griffin, W.; Swendsen, R.H. Comparison of canonical and microcanonical definitions of entropy. *Physica A* **2017**, *467*, 474–489.
24. Buonsante, P.; Franzosi, R.; Smerzi, A. On the dispute between Boltzmann and Gibbs entropy. *Ann. Phys.* **2016**, *375*, 414–434.
25. Buonsante, P.; Franzosi, R.; Smerzi, A. Phase transitions at high energy vindicate negative microcanonical temperature. *Phys. Rev. E* **2017**, *95*, 052135.
26. Swendsen, R.H. The definition of the thermodynamic entropy in statistical mechanics. *Physica A* **2017**, *467*, 67–73.
27. Abraham, E.; Penrose, O. Physics of negative absolute temperatures. *Phys. Rev. E* **2017**, *95*, 012125.
28. Purcell, E.M.; Pound, R.V. A nuclear spin system at negative temperature. *Phys. Rev.* **1951**, *81*, 279–280.
29. Ramsey, N.F. Thermodynamics and Statistical Mechanics at Negative Absolute Temperatures. *Phys. Rev.* **1956**, *103*, 20–28.
30. Tisza, T. *Generalized Thermodynamics*; MIT Press: Cambridge, MA, USA, 1966.
31. Callen, H.B. *Thermodynamics*; Wiley: New York, NY, USA, 1960.
32. Callen, H.B. *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed.; Wiley: New York, NY, USA, 1985.
33. Swendsen, R.H. Explaining Irreversibility. *Am. J. Phys.* **2008**, *76*, 643–648.
34. Swendsen, R.H. Irreversibility and the Thermodynamic Limit. *J. Stat. Phys.* **1974**, *10*, 175–177.
35. Boltzmann, L. Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht. *Wien. Ber.* **1877**, *76*, 373–435; reprinted in *Wissenschaftliche Abhandlungen von Ludwig Boltzmann*; Chelsea: New York, NY, USA; Volume II, pp. 164–223.
36. Boltzmann derived the expression for the entropy for classical systems under the assumption that equilibrium corresponded to the maximum of the probability for two systems in equilibrium. The quantum version is due to Planck. Planck is also responsible for the form of the entropy carved into Boltzmann’s grave stone, $S = k \log W$ (where “W” stands for the German word “Wahrscheinlichkeit”, or probability), and the introduction of k as Boltzmann’s constant.
37. Sharp, K.; Matschinsky, F. Translation of Ludwig Boltzmann’s Paper On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium, *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften. Mathematisch-Naturwissen Classe. Abt. II, LXXVI* 1877, pp. 373–435 (Wien. Ber. 1877, 76:373–435). Reprinted in *Wiss. Abhandlungen, Vol. II, reprint 42*, pp. 164–223, Barth, Leipzig, 1909. *Entropy* **2015**, *17*, 1971–2009.
38. Planck, M. Über das Gesetz der Energieverteilung im Normalspektrum. *Drudes Annalen* **1901**, *553*, 65–74; reprinted in *Ostwalds Klassiker der exakten Wissenschaften, Band 206, “Die Ableitung der Strahlungsgesetze”*.
39. Planck, M. *Theorie der Wärmestrahlung*; Barth Verlag: Leipzig, Germany, 1906. Translated by M. Masius and reprinted as *The Theory of Heat Radiation*; Dover: New York, NY, USA, 1991.
40. Lee, J.; Kosterlitz, J.M. New Numerical Method to Study Phase Transitions. *Phys. Rev. Lett.* **1990**, *65*, 137–140.
41. Berg, B.A.; Neuhaus, T. Multicanonical Ensemble: A New Approach to Simulate First-Order Phase Transitions. *Phys. Rev. Lett.* **1992**, *68*, 9–12.
42. Hüller, A. First order phase transitions in the canonical and the microcanonical ensemble. *Zeitschrift für Physik B* **1994**, *93*, 401–405.

43. Swendsen, R.H. *An Introduction to Statistical Mechanics and Thermodynamics*; Oxford University Press: London, UK, 2012.
44. Griffin, W.; Matty, M.; Swendsen, R.H. Finite thermal reservoirs and the canonical distribution. *Physica A* **2017**, *484*, 1–10.
45. Loeb, E.M. The third law of thermodynamics, the unattainability of absolute zero, and quantum mechanics. *J. Chem. Educ.* **1960**, *37*, 361–363.
46. Fowler, R.; Guggenheim, E.A. *Statistical Thermodynamics. A Version of Statistical Mechanics for Students of Physics and Chemistry*; Cambridge University Press: Cambridge, UK, 1939.
47. I am distinguishing extensivity from additivity. The entropy of a system is extensive when $\lambda S(U, V, N) = S(\lambda U, \lambda V, \lambda N)$. The entropies of two systems are additive when $S_{A,B} = S_A + S_B$.
48. Swendsen, R.H. The ambiguity of ‘distinguishability’ in statistical mechanics. *Am. J. Phys.* **2015**, *83*, 545–554.
49. If any two systems can exchange both volume and particles through a piston with a hole in it, a range of positions of the piston will have the same value of W .
50. If a system has N particles the fluctuations will be of order $1/\sqrt{N}$, and it will require of the order of N independent measurements to determine a difference of order $1/N$. For $N = 10^{12}$ particles, if a measurement of perfect accuracy were to be made every second, it would take over 30,000 years to detect difference of order $1/N$. For $N = 10^{20}$, it would take about 200 times the age of the universe.
51. Gibbs, J.W. *Elementary Principles of Statistical Mechanics*; Yale University Press: New Haven, CT, USA, 1902; reprinted by Dover, New York, 1960.
52. I am using the term “Gibbs entropy” to refer to the definition of the entropy in terms of the logarithm of the volume of phase space with energy less than a given energy. The quantum version refers to the sum of all degeneracies of eigenstates with energies below a given energy. It is not to be confused with another definition of entropy due to Gibbs in terms of an integral of $\int \rho \ln \rho$, where ρ is the probability of a microscopic state.
53. Hertz, P. Über die mechanischen Grundlagen der Thermodynamik. *Ann. Phys. (Leipz.)* **1910**, *338*, 225–274.
54. Khinchin, A.I. *Mathematical Foundations of Statistical Mechanics*; Dover: New York, NY, USA, 1949.
55. O. Penrose (Heriott-Watt University, Edinburgh, Scotland, UK) and J.-S. Wang (National University of Singapore, Singapore, Singapore) have both separately raised this point. Private communications (2015).
56. Touchette, H.; Ellis, R.S.; Turkington, B. An introduction to the thermodynamic and macrostate levels of nonequivalent ensembles. *Physica A* **2004**, *340*, 138–146.
57. Touchette, H. The large deviation approach to statistical mechanics. *Phys. Rep.* **2009**, *478*, 1–69.
58. Touchette, H. Methods for calculating nonconcave entropies. *J. Stat. Phys.* **2010**, *P05008*, 1–22.
59. Touchette, H. Ensemble equivalence for general many-body systems. *Europhysics Letters* **2011**, *96*, 50010.
60. Touchette, H. Equivalence and nonequivalence of ensembles: Thermodynamic, macrostate, and measure levels. *J. Stat. Phys.* **2015**, *159*, 987–1016.

