Unnormalized probability: A different view of statistical mechanics

Robert H. Swendsen^{a)}

Physics Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

(Received 24 July 2013; accepted 3 June 2014)

All teachers and students of physics have absorbed the doctrine that probability must be normalized. Nevertheless, there are problems for which the normalization factor only gets in the way. An important example of this counter-intuitive assertion is provided by the derivation of the thermodynamic entropy from the principles of statistical mechanics. Unnormalized probabilities provide a surprisingly effective teaching tool that can make it easier to explain to students the essential concept of entropy. The elimination of the normalization factor offers simpler equations for thermodynamic equilibrium in statistical mechanics, which then lead naturally to a new and simpler definition of the entropy in thermodynamics. Notably, this definition does not change the formal expression of the entropy based on composite systems that I have previously offered. My previous definition of entropy has been criticized by Dieks, based on what appears to be a misinterpretation. I believe that the new definition presented here has the advantage of greatly reducing the possibility of such a misunderstanding—either by students or by experts. © 2014 American Association of Physics Teachers. [<http://dx.doi.org/10.1119/1.4883480>]

I. INTRODUCTION

An important goal in teaching thermal physics is to explain to students how statistical mechanics provides a microscopic foundation for the thermodynamics of macroscopic systems. The explanation should take the students from the probability distribution for the microscopic states to the thermodynamic entropy with as few assumptions as possible and no arbitrary definitions to take on faith.

With this goal in mind, I propose an alternative way of looking at the mathematical structure of statistical mechanics and its relationship to thermodynamic properties. This alternative view begins with the usual idea that limited experimental information of the microscopic state of a macroscopic (many-particle) system leads to the use of probabilities to describe available knowledge. The principle that all microscopic states consistent with available information about the system are equally likely allows the calculation of probability distributions for macroscopic observables in isolated composite systems. The probability distributions for these macroscopic observables are predicted to have very small statistical uncertainties.

An important property of probability is that it is normalized—its sum (or integral) over all microscopic states is one. Indeed, the partition function, which is a common basis for calculations in statistical mechanics, is essentially a normalization factor. However, the entropy is related to the logarithm of the probability,^{1,2} so a normalization constant would appear in the equations for the entropy as an irrelevant additive constant.^{3,4} Expressing equilibrium conditions in terms of unnormalized probabilities eliminates this irrelevant constant before considering the connection between statistical mechanics and the thermodynamic entropy. As I will show below, the result is a clear derivation of the thermodynamic entropy as a function of energy, volume, and particle number from statistical mechanics. Since the entropy as a function of the extensive variables is a fundamental relation, all thermodynamic properties can then be derived without any need for additional calculations in statistical mechanics.⁵

Most of the discussion in this paper will be in terms of classical statistical mechanics. The extension to quantum mechanics is straightforward and will be presented briefly in Sec. VII.

I will begin by briefly reviewing normalized probabilities in statistical mechanics and then discuss the advantages of discarding normalization factors to calculate equilibrium values of macroscopic observables. This will provide a framework for the new definition of the thermodynamic entropy, which will be given in subsequent sections.

II. THERMAL EQUILIBRIUM IN CLASSICAL STATISTICAL MECHANICS

A central problem in thermal physics is the prediction of the values of macroscopic observables after the release of a constraint. This problem is most easily analyzed by considering composite systems, in which there is an initial partitioning of the particles and/or the energy. The composite system can then be regarded as being composed of simple systems, each of which has an entropy that is to be calculated from statistical mechanics. The entropy of an arbitrary simple system of interest will be a function of its macroscopic variables: energy E , volume V , and number of particles N .⁶

Suppose the system of interest is brought into contact with another system—also arbitrary—and characterized by the macroscopic variables E' , V' , and N' . We want to calculate what will happen when this composite system is isolated from the rest of the universe and one or more constraints are released. For example, two systems might be put in thermal contact with each other so that they can exchange energy, and we would like to predict the final values of E and E' when the system reaches equilibrium.

As mentioned in the Introduction, a solution to this problem is provided by a normalized probability distribution determined by the condition that all microscopic states consistent with the known macroscopic values are equally likely. I'll review that solution in Sec. IV, but first I'll illustrate the alternative approach using unnormalized probabilities with a simple example in Sec. III.

The full solution in terms of unnormalized probabilities, including the extension to exchanges between arbitrarily many systems, will be given in Sec. V. The implications of this solution for the definition of the thermodynamic entropy of arbitrarily many systems will be given in Sec. VI.

III. A SIMPLE EXAMPLE OF UNNORMALIZED PROBABILITIES

The simplest example to illustrate the advantages of unnormalized probabilities is the distribution of classical ideal-gas particles divided between two volumes V and V' . The composite system is isolated from the rest of the universe. The initial numbers of particles in the two subsystems are N_0 and N'_0 , and a hole is made in the partition between the two containers. The total number of particles $N_T = N_0 + N'_0 = N + N'$ is constant because the composite system is isolated. The probability of finding N particles in a volume V is given by the binomial distribution

$$W_{\text{binom}}(N, N') = \frac{N_T!}{N!N'} \frac{V^N V'^{N'}}{V_T^{N_T}}, \quad (1)$$

where $V_T = V + V'$.^{5,7} In this equation, W_{binom} has been written as a function of both N and N' to highlight the symmetry between the two systems, but it is to be evaluated with $N' = N_0 + N'_0 - N = N_T - N$.

It is well known that the binomial distribution for $N/(N_0 + N'_0) = N/N_T$, the fraction of particles in volume V is sharply peaked for large numbers of particles. With enough particles, the width of the probability distribution will be smaller than the experimental resolution, and we can take the location of the maximum of $W_{\text{binom}}(N, N_0 + N'_0 - N)$ as the equilibrium value of N .

Now consider the unnormalized variant of Eq. (1):

$$\hat{W}_{\text{binom}}(N, N') = \frac{V^N V'^{N'}}{N! N'!}. \quad (2)$$

I have used a hat to indicate that $\hat{W}_{\text{binom}}$ is not normalized. Since $\hat{W}_{\text{binom}}$ has its maximum at exactly the same value of N as W_{binom} , the location of its maximum also predicts the equilibrium value of N . The normalization factor $N_T! / V_T^{N_T}$ has no effect.

The unnormalized probability $\hat{W}_{\text{binom}}$ is somewhat easier to work with than W_{binom} . For one, it has a simple form that consists of a product of two factors, $V^N / N!$ and $V'^{N'} / N'!$, each of which depends only on the properties of one of the subsystems. In Sec. IV, I will show that there are other advantages.

It should be noted that there is an ambiguity in defining $\hat{W}_{\text{binom}}$. Because

$$\frac{A^N A'^{N'}}{A^{N_T}} = 1 \quad (3)$$

for any constant value of A , we could also define the factor for the unprimed system as $A^N V^N / N!$, with a comparable expression for the primed system. For this simple example, there is no reason to choose any value of A other than one. Nevertheless, Sec. IV shows that the flexibility of being able to choose A can be convenient in dealing with general macroscopic systems.

IV. NORMALIZED PROBABILITIES AND THERMAL EQUILIBRIUM

We now extend the discussion in Sec. III to the more general problem of equilibrium between two macroscopic systems with respect to energy, volume, or particle number.

Consider two classical systems with arbitrary short-ranged interactions between the particles. Define the set of all momenta as $\{p\}$ and the set of all position coordinates as $\{q\}$, so that a point in the joint set $\{p, q\}$ represents a microscopic state of the system. The Hamiltonians will have the form

$$H(p, q) = \sum_{j=1}^N \frac{|\vec{p}_j|^2}{2m} + \sum_{j>k} \phi(\vec{r}_j, \vec{r}_k), \quad (4)$$

where ϕ is the interaction energy for a pair of particles.

The assumption of a uniform probability distribution in phase space is subject to the constraint that the particles are in volumes V and V' . The energies of the two systems are specified by delta functions $\delta(E - H(p, q))$ and $\delta(E' - H'(p', q'))$ for the unprimed and primed systems. These assumptions lead to a probability function of the form⁵

$$W_2(E, V, N; E', V', N') = \frac{\Omega(E, V, N) \Omega'(E', V', N')}{\Omega_T}. \quad (5)$$

In this equation, Ω depends only on the properties of the unprimed system, and Ω' depends only on the properties of the primed system; Ω_T is a normalization factor. As shown elsewhere,⁵ the functions Ω and Ω' in Eq. (5) can be expressed as

$$\Omega(E, V, N) = \frac{1}{h^{3N}} \frac{1}{N!} \int dp \int dq \delta(E - H(p, q)). \quad (6)$$

In this equation, I have used the freedom to multiply Ω by a factor of A^N , as explained in Sec. III, to choose the universal value $A = 1/h^3$ for a three-dimensional system. Since h is intended to be Planck's constant, it is clear that this factor has nothing to do with classical statistical mechanics. However, it is convenient to include the factor of $1/h^{3N}$ because it ensures that the classical free energy agrees with the classical limit of the quantum free energy. A demonstration of this is beyond the scope of this paper, but can be found elsewhere.⁵ The factor of $1/N!$ has the same origin as the corresponding factor in Eqs. (1) and (2). For an ideal gas, the integral over the positions would give the factor V^N , which is also found in Eqs. (1) and (2).

Note that Eq. (6) is not the multiplicity of the states in an isolated system, nor is it a volume in phase space. The presence of the factor $1/N!$ arises only because of its derivation in the context of a composite system.

We isolate our two systems (unprimed and primed) from the rest of the universe, and bring them together to form a composite system, as described in Sec. II. Now we can calculate the equilibrium values of the macroscopic observables when a constraint is released. For example, we can put the two systems into thermal contact so they can exchange energy, while keeping the total energy $E_T = E_0 + E'_0 = E + E'$ constant. As expected, we find that the probability distribution W_2 for E/E_T is very sharply peaked, so that we can take the location of the peak as the equilibrium value.

For the determination of the location of the equilibrium values of thermodynamic quantities, the factor Ω_T in Eq. (5) is irrelevant. Its magnitude will affect the height of the peak in W_2 , but it cannot affect the location of the peak.

In addition to being irrelevant, the factor Ω_T is ambiguous. In describing Eq. (5), I referred to Ω_T in the denominator as

a normalization factor, not *the* normalization factor. The reason is that the normalization of W_2 depends on which constraints are to be released and which constraints retained. For example, if we let the systems exchange energy, but the numbers of particles in each system are fixed, the proper normalization is to divide $\Omega\Omega'$ by

$$\int_0^{E_T} \Omega(E, V, N) \Omega'(E_T - E, V', N') dE, \quad (7)$$

while keeping V, N, V' , and N' constant; Ω_T would be a function of E_T, V, N, V' , and N' .

On the other hand, if we poke a hole in the partition between the systems, both energy and particles could be exchanged. The proper normalization is then to divide $\Omega\Omega'$ by

$$\sum_{N=0}^{N_T} \int_0^{E_T} \Omega(E, V, N) \Omega'(E_T - E, V', N_T - N) dE, \quad (8)$$

while keeping V and V' constant. Now Ω_T would be a function of E_T, V, V' , and N_T . The normalization constants in these two examples are quite different and even depend on different variables.

The normalization constants in Eqs. (7) and (8) have no effect on the determination of the equilibrium values of E and N from the maximum of W . The consequences of this irrelevance will be explored in Sec. V.

V. UNNORMALIZED PROBABILITIES

Since Sec. IV showed that a normalization factor in the denominator of W_2 in Eq. (5) is both ambiguous and irrelevant to the determination of equilibrium values, we can effectively and, as it turns out, advantageously discard it. Denote the unnormalized probability distribution for the energy, volume, and number of particles in the composite system by giving it a hat, so that

$$\hat{W}_2(E, V, N; E', V', N') = \Omega(E, V, N) \Omega'(E', V', N'). \quad (9)$$

The values of the macroscopic observables in equilibrium for any choice of imposed or released constraints in the composite system are given by the location of the maximum of $\hat{W}_2$.

Note that the factors $\Omega(E, V, N)$ and $\Omega'(E', V', N')$ are not the volumes in phase space for the subsystems since, as shown in Eq. (6), the integrals contain a delta function that selects a surface of constant energy. The factors are also not multiplicities of the states in each of the subsystems due to the factors of $1/N!$ and $1/N'!$, which come from considering a composite system.

The clarity obtained by using unnormalized probabilities becomes particularly evident when we consider equilibrium between three systems. The unnormalized probability $\hat{W}_3$ is then just the product of three factors, each of which only depends on the properties of a single system:

$$\hat{W}_3 = \Omega(E, V, N) \Omega'(E', V', N') \Omega''(E'', V'', N''). \quad (10)$$

The function $\hat{W}_3$ will exhibit a maximum at the equilibrium value of any of its arguments when any desired constraints are imposed or released.

The ambiguity of a “normalization constant” for three systems is much greater than for two. We could use $\hat{W}_3$ to find the equilibrium value of E under four conditions: (1) all three systems in thermal contact, (2) E held constant, (3) E' held constant, or (4) E'' held constant. The appropriate normalization constant is different in each of these four cases. We might also poke holes in the walls between different systems to produce four more normalization constants with distinct values. Further possibilities are obvious.

The conclusion is that the single function $\hat{W}_n$ is completely sufficient to determine the equilibrium conditions for any experiment involving exchanges of energy, volume, or particles between n subsystems; using the normalized probability W_n would require many different functions, each with a different normalization constant. Using the unnormalized probability therefore simplifies the calculation of equilibrium conditions in statistical mechanics without losing any information.

So far, I have only shown that the calculation of equilibrium conditions in statistical mechanics can be simplified by the use of unnormalized probabilities. In Sec. VI, I will use these results to give a simple definition of the thermodynamic entropy from statistical mechanics.

VI. DERIVATION OF ENTROPY FROM UNNORMALIZED PROBABILITIES

Now consider the fundamental question of how best to define the thermodynamic entropy of a system of interest from statistical mechanics. The most important property of the thermodynamic entropy is that it is maximized in equilibrium. More precisely, when a constraint is released in an isolated composite system, the total entropy goes to a maximum. This is the essence of the second law of thermodynamics: changes of entropy in an isolated system are never negative.

We are also looking for a definition of entropy that has the property of additivity for systems with short-range interactions between particles. That is, we want to be able to identify the entropy of any macroscopic system, and construct the total entropy of a composite system by simply forming the sum of the individual entropies.

In Sec. V, we found that the function $\hat{W}_2$ in Eq. (9) has the desired property that its maximum always gives the location of the equilibrium values of E, V , or N . However, it is not a suitable candidate for entropy because the contributions of the two systems are in the form of a product: $\hat{W}_2 = \Omega\Omega'$. The natural thing to do is to follow Boltzmann in taking the logarithm of $\hat{W}_2$ to define the entropy:^{1,2}

$$S_2 = k_B \ln \hat{W}_2(E, V, N; E', V', N'). \quad (11)$$

While Boltzmann never used the “Boltzmann constant” k_B , I have followed Planck by including k_B in the definition of the entropy to ensure that the units are correct.⁸ Now, because

$$\ln \hat{W}_2 = \ln \Omega(E, V, N) + \ln \Omega'(E', V', N'), \quad (12)$$

the entropy of an individual system is given by

$$S(E, V, N) = k_B \ln \Omega(E, V, N), \quad (13)$$

so that

$$S_2(E, V, N; E', V', N') = S(E, V, N) + S'(E', V', N'), \quad (14)$$

with the explicit conditions that the total energy ($E_T = E + E'$), total volume ($V_T = V + V'$), and total number of particles ($N_T = N + N'$) are constant.

The final extension of the derivation of the entropy to general composite systems is straightforward. We simply carry out the same analysis for a set of n subsystems, where n can be arbitrarily large. Denoting the probability factor of the j -th system as $\Omega^{(j)}$, the unnormalized probability distribution for n subsystems is given by

$$\hat{W}_n = \prod_{j=1}^n \Omega^{(j)}(E^{(j)}, V^{(j)}, N^{(j)}), \quad (15)$$

with only the conditions that $E_{n,\text{total}} = \sum_{j=1}^n E^{(j)}$, $V_{n,\text{total}} = \sum_{j=1}^n V^{(j)}$, and $N_{n,\text{total}} = \sum_{j=1}^n N^{(j)}$ are constant.

For any experiment with any constraints imposed on any combination of the variables $\{E^{(j)}, V^{(j)}, N^{(j)} | j = 1, \dots, n\}$, the location of the maximum of W_n will give the equilibrium condition. This means that the equilibrium conditions are also given by the maximum of the total entropy

$$S_n = \sum_{j=1}^n S^{(j)}(E^{(j)}, V^{(j)}, N^{(j)}), \quad (16)$$

where

$$S^{(j)}(E^{(j)}, V^{(j)}, N^{(j)}) = k_B \ln \Omega^{(j)}(E^{(j)}, V^{(j)}, N^{(j)}). \quad (17)$$

This demonstrates that the equilibrium conditions for an arbitrary composite system in contact with an arbitrary set of other systems are also determined by the maximum of the sum of the entropies of the individual systems. This releases, for example, the constraints on the total energy, volume, and particle number of a composite system, so that the expression for the entropy given in Eq. (14) is generally valid.

Since the functional dependence of the entropy on the energy, volume, and particle number is a fundamental relation in thermodynamics, all thermodynamic information can be obtained from this function. Inversion of this function gives the same information in the energy representation $U = U(S, V, N)$. Legendre transforms give the fundamental relation in other representations, such as the Helmholtz free energy $F = F(T, V, N)$ and the Gibbs free energy $G = G(T, P, N)$.⁵

This completes the derivation of the entropy for classical many-particle systems. The implications of the unnormalized approach for the entropy of quantum many-particle systems are discussed in Sec. VII.

VII. QUANTUM STATISTICAL MECHANICS AND PLANCK'S "THERMODYNAMIC PROBABILITIES"

The most commonly used expression for the entropy of a quantum system is due to Planck;^{8,9} it can be written as

$$S_{QM}(E, V, N) = k_B \ln \Omega_{QM}(E, V, N), \quad (18)$$

where $\Omega_{QM}(E, V, N)$ is the degeneracy of the quantum energy level E . Planck called $\Omega_{QM}(E, V, N)$ the "thermodynamic probability," but many have objected to that term because $\Omega_{QM}(E, V, N)$ is neither normalized nor normalizable.

On the other hand, if we have an isolated composite system, the product

$$\hat{W}_{2,QM} = \Omega_{QM}(E, V, N) \Omega'_{QM}(E', V', N') \quad (19)$$

is the unnormalized probability for the energy distribution when the total energy is held constant.¹⁰ It is therefore consistent with the discussion given above for classical systems to define the entropy of a composite quantum system as

$$S_{2,QM}(E, V, N; E', V', N') = k_B \ln \hat{W}_{2,QM}, \quad (20)$$

which gives

$$S_{2,QM} = S_{QM}(E, V, N) + S'_{QM}(E', V', N'), \quad (21)$$

where the entropy of an individual system is given by Eq. (18). While $\Omega_{QM}(E, V, N)$ is not a probability, it is a factor in the unnormalized probability $\hat{W}_{2,QM}(E, V, N; E', V', N')$ in Eq. (19). Unnormalized probabilities of composite quantum systems therefore also give the correct expression for the entropy from quantum statistical mechanics.

Although Boltzmann never to my knowledge expressed his ideas in terms of unnormalized probabilities, I believe that Eqs. (11) and (20), $S_2 = k_B \ln \hat{W}_2$, are closer to Boltzmann's intention than the equation $S = k \log W$, which is displayed on Boltzmann's monument in the Wiener Zentralfriedhof. Equations (11) and (20) explicitly define the entropy in terms of a composite system, as did Boltzmann,^{1,2} and W_2 is a probability ("Wahrscheinlichkeit" in German), although admittedly not normalized.

VIII. OTHER DEFINITIONS OF ENTROPY

The most common way to teach entropy is to present an equation based on the properties of an isolated system. This relies on an argument to authority, which I believe to be a poor way to teach science. It requires claiming that an equation is true because Boltzmann said so, or Gibbs said so, or the textbook said so and this can give students the impression that learning science is memorizing equations with obscure justifications. That is the wrong message.

The isolated-system approach has also led to a great deal of confusion for more than a century about the dependence of the entropy on the number of particles. About thirty years ago, van Kampen noted that

The dependence of the entropy on the number of molecules can never be found from studying closed systems. The argument based on counting states of particles, indistinguishable or not, is therefore spurious.¹¹

Defining the entropy through the properties of composite systems, as I suggested in 2002,³—and Boltzmann 125 years earlier²—satisfies van Kampen's criterion of being based on the behavior of systems that can exchange particles (and energy, and volume) with another system. This definition always leads to the correct thermodynamic predictions for all possible macroscopic experiments, and Eq. (13) provides a correct fundamental relation for an arbitrary macroscopic system. Since the fundamental relation is unique, any alternative definition of the thermodynamic entropy must be equivalent to this one.

IX. MY EARLIER DEFINITION OF ENTROPY

In 2002, I suggested a definition of entropy that differed from the one presented here,³ although it was also based on an analysis of composite systems and turns out to be equivalent. In the earlier definition I began with the normalized probability, which led to the following equation for two simple systems:⁴

$$k_B \ln W_2 + k_B \ln \Omega_T = k_B \ln \Omega + k_B \ln \Omega'. \quad (22)$$

The expression on the left side of the equation must be a maximum at equilibrium, under the usual conditions that $E_2 = E + E'$, $V_2 = V + V'$, and $N_2 = N + N'$ are all constant. Therefore, the right side must be also be a maximum at equilibrium under the same conditions. However, the right side of Eq. (22) is a simple sum of two functions, each of which depends only on variables in one of the subsystems. Since this is exactly the condition specified by the second law of thermodynamics, the two expressions on the right side can be identified as the entropies of the subsystems. This gave exactly the same expression for $S(E, V, N)$ as in Eq. (13).

As discussed in Sec. VI, the dependence of the entropy of a composite system on E , V , and N can only be determined from the analysis of exchanges with other systems as part of a larger composite system.⁴

X. THE OBJECTION OF DIEKS

My earlier definition of the entropy using composite systems has been criticized by Dieks,¹² who claimed that by eliminating the additive constant $-k_B \ln \Omega_T$, I had arbitrarily introduced a term of the form $-k_B \ln N!$. He concluded that my definition of the entropy is not “a fundamental microscopic justification of the division by $N!$.” I believe Dieks’ criticism is the result of a misrepresentation of my definition of entropy, and his key error is to treat his constant N as if it were a variable.

Dieks begins his argument by defining a composite system consisting of two subsystems with N_1 and N_2 particles, respectively, and defines $N = N_1 + N_2$ as the total number of particles in the composite system; N is constant because the composite system is isolated. He restricts his discussion to a classical ideal gas and ignores the energy dependence, apparently because he has no objections to my treatment of it. He then follows my argument to show in his Eq. (2) that the entropy terms for the two subsystems contain factors of $1/N_1!$ and $1/N_2!$ as I had claimed. This seems to show that Dieks agrees with my derivation of these factors, as does his later comment that, “the dependence of the total entropy in Eq. (2) on N_1 and N_2 is unrelated to how N occurs in this formula.”¹²

A reasonable conclusion from Dieks’ derivation and comments is that he has accepted my argument for the presence of the factors of $1/N_1!$ and $1/N_2!$ in the entropies of his subsystems 1 and 2, and therefore in the entropy of any classical ideal gas. His formulation further suggests his acceptance of my Eqs. (6) and (13) for the general expression for the entropy (where I had used N in Sec. VI for the number of particles in a simple system). However, Eq. (6) is the only place that a factor of $1/N!$ (or $1/N_j!$, or $1/N^{(j)}!$) enters my definition of the entropy. It would seem that Dieks has already disproven his own objection.

At this point, Dieks turns to the question of a possible additive constant in the entropy of a composite system. Since his composite system is isolated, the total volume and particle number are constant, and there could be an additive constant that depends on V and N . I had included the calculation of this constant in my definition of the entropy, and pointed out that, “Since the entropy of any thermodynamic system can be calculated by the procedure described above [referring to my procedure to derive the entropy of a simple system] we can also use it to find the entropy of a composite system.”⁴ An explicit derivation of the entropy of composite systems, which is essentially the same as for my older definition, is contained in Sec. VI of this paper.

If Dieks had followed my definition, he would have found that Eq. (14) also gives the appropriate expression for the entropy of a composite system that can exchange energy, volume, or particles with other systems. There is no extra additive constant.

However, Dieks does not follow my definition, but rather writes that it fixed the additive constant “by adding the freely chosen constant $k \ln V^N/N!$ to the logarithm of the probability.”¹² This would have been correct for the special case that Dieks now considers if he had realized that it was a constant and could not be interpreted as a variable, as he proceeds to do.

Dieks then considers the special case in which “the two subsystems can exchange particles,” from which he obtains his Eq. (4), which is repeated here:

$$k_B \ln \frac{V_1^{N_1}}{N_1!} + k_B \ln \frac{V_2^{N_2}}{N_2!} \approx k_B \ln \frac{V^N}{N!}. \quad (23)$$

Dieks has converted his composite system into a simple system consisting of an isolated ideal gas of N particles in a volume V .

Dieks then writes: “Swendsen claims that in this way the factor $1/N!$ in the formula for the entropy has been demonstrated to be a necessary consequence of the distinguishability of the gas atoms or molecules.”¹² This is the crucial misrepresentation of my definition. I had derived the dependence of the entropy of the subsystems on the variables N_1 and N_2 . Dieks’ system is still isolated so that N and V are constants, not variables. I have consistently maintained, as did van Kampen,¹¹ that the N -dependence of the entropy cannot be determined by any argument involving only an isolated system. Since Dieks’ entire objection to my definition appears to rest on his misunderstanding of the significance of the additive constant—and, indeed, his forgetting the fact that it is a constant—I maintain that his objection is invalid.

I believe that the modified definition presented in this paper makes it easier to avoid the error above. The normalization constant Ω_T has already been eliminated from the formulation of equilibrium conditions in statistical mechanics in terms of unnormalized probabilities, so it never enters the new definition of entropy. Since Ω_T is not there, there should be no temptation to misinterpret its significance.

One final point should be mentioned. Dieks concentrates on the extensivity of the entropy of a homogeneous system: a classical ideal gas. However, my definition is also applicable to intrinsically non-extensive systems like a real gas that can be adsorbed onto the walls of its container.

XI. THE CONSEQUENCES OF DIFFERENT DEFINITIONS OF ENTROPY

In comparing Dieks' views on entropy with my own, it can be seen that they have very different consequences. My goal is to use statistical mechanics to obtain a fundamental relation in the form of the thermodynamic entropy. That is, my definition gives a function $S = S(E, V, N)$ that contains all thermodynamic information. There is no need to return to statistical mechanics to make additional calculations for that system.

The consequences of the calculations by Dieks and his collaborator Versteegh are very different.^{12,13} Their approach results in a separate calculation in statistical mechanics for every combination of interacting systems. This becomes clear from their response to my statement that the traditional expression for the entropy of a classical ideal gas found without the factor of $1/N!$,

$$S_{\text{trad}}(E, V, N) = k_B N \left[\frac{3}{2} \ln \left(\frac{E}{N} \right) + \ln V + \text{constant} \right], \quad (24)$$

gives incorrect predictions for a composite system.³ Rather than defending the expression in Eq. (24) that I had criticized, they resorted to a new statistical mechanics calculation for the composite system.¹³ My objection remains valid.

I believe that our difference in approach is critically important. As I have pointed out earlier, "the concepts of entropy, free energy, etc. are extremely convenient, but they are not absolutely necessary. We could calculate anything and everything about the behavior of macroscopic systems without ever mentioning them."¹⁴ The drawback is that you would have to do a new calculation in statistical mechanics for every combination of systems; you would not be able to use the equations of thermodynamics.

The great advantage of obtaining a fundamental relation like $S = S(E, V, N)$ is that it contains all information about the system needed to calculate the results of any thermodynamic interaction with any other system. No further recourse to statistical mechanics is necessary.

XII. CONCLUSIONS

I believe that the derivation of the thermodynamic entropy from unnormalized probabilities in statistical mechanics

provides the clearest way to teach upper-level students about this central concept. The calculation of equilibrium values for macroscopic variables from unnormalized probabilities in composite systems clarifies the connection between statistical mechanics and thermodynamics and aids students in understanding the origin and meaning of the thermodynamic entropy.

^{a)}Electronic mail: swendsen@cmu.edu

¹R. H. Swendsen, "Statistical mechanics of colloids and Boltzmann's definition of the entropy," *Am. J. Phys.* **74**, 187–190 (2006).

²L. Boltzmann, "Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht," *Wien. Ber.* **76**, 373–435 (1877), reprinted in *Wissenschaftliche Abhandlungen von Ludwig Boltzmann* (Chelsea, New York, 1968), Vol. II, pp. 164–223.

³R. H. Swendsen, "Statistical mechanics of classical systems with distinguishable particles," *J. Stat. Phys.* **107**, 1143–1165 (2002).

⁴R. H. Swendsen, "Choosing a definition of entropy that works," *Found. Phys.* **42**, 582–593 (2012).

⁵R. H. Swendsen, *An Introduction to Statistical Mechanics and Thermodynamics* (Oxford, London, 2012).

⁶I'll limit the discussion to systems with a single type of particle. Although it is straightforward to extend the formalism to many types of particles, it is not essential to the main argument and would only complicate the notation.

⁷Some students, and even some professors, accept Eq. (1) for states that are not particle exchange symmetric but not for states that are. It is true for both.

⁸M. Planck, "Über das Gesetz der Energieverteilung im Normalspektrum," *Drudes Ann.* 553–562 (1901), reprinted in *Ostwalds Klassiker der exakten Wissenschaften, Band 206*, "Die Ableitung der Strahlungsgesetze," pp. 65–74.

⁹M. Planck, *Theorie der Wärmestrahlung* (J. A. Barth, Leipzig, 1906), translated into English by Morton Masius in M. Planck, *The Theory of Heat Radiation* (Dover, New York, 1991), p. 119.

¹⁰There is a subtlety involved in applying Eq. (19) because the spectra of the primed and unprimed systems do not generally allow arbitrary changes in E and E' when the sum $E_T = E + E'$ is held constant. This mathematical difficulty is most commonly dealt with by letting one of the systems become infinitely large and using it to define the canonical ensemble.⁵

¹¹N. G. van Kampen, "The Gibbs paradox," in *Essays in Theoretical Physics*, edited by W. E. Parry (Pergamon, Oxford, 1984).

¹²D. Dieks, "Is There a Unique Physical Entropy? Micro versus Macro," in *New Challenges to Philosophy of Science* (The Philosophy of Science in a European Perspective), edited by H. Andersen, D. Dieks, W. Gonzalez, T. Uebel, and G. Wheeler (Springer, New York, 2013).

¹³M. A. M. Versteegh and D. Dieks, "The Gibbs paradox and the distinguishability of identical particles," *Am. J. Phys.* **79**, 741–746 (2011).

¹⁴R. H. Swendsen, "How physicists disagree on the meaning of entropy," *Am. J. Phys.* **79**, 342–348 (2011).