

Statistical mechanics of colloids and Boltzmann's definition of the entropy

Robert H. Swendsen

Physics Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

(Received 17 November 2005; accepted 29 December 2005)

The Boltzmann entropy as traditionally presented in statistical mechanics textbooks is only a special case and not Boltzmann's fundamental definition. The difference becomes important when the traditional expression for the entropy is applied to colloids, for which it makes incorrect predictions. Boltzmann's original definition of the entropy in terms of the probabilities of states of composite systems leads to consistent and correct statistical mechanics and thermodynamics. © 2006 American Association of Physics Teachers.
[DOI: 10.1119/1.2174962]

I. INTRODUCTION

Colloids represent a problem for traditional statistical mechanics. They should be well represented by the results of classical statistical mechanics for distinguishable particles. Unfortunately, the usual expressions make nonsensical predictions.

In this paper I will discuss the problem with the traditional textbook definition of entropy, trace it to a misunderstanding of Boltzmann's work, and show how the direct application of Boltzmann's fundamental definition of the entropy eliminates the difficulties.

Colloids consist of a large number of small particles of one material distributed uniformly in a different material.¹ There are many different kinds of colloids, including paint, ink, foam, glue, blood, and homogenized milk, and they are of great commercial importance. The size of the suspended particles varies from about a nanometer to several microns. Colloidal particles are usually sufficiently large to be well described by classical mechanics. They also differ from one another significantly because they usually contain different numbers of atoms in different arrangements with different impurities. These properties makes them distinguishable in the usual sense that the exchange of two particles produces a different microscopic state. For these reasons colloids should be well described by the classical statistical mechanics of distinguishable particles.

II. THE TRADITIONAL CALCULATION OF THE ENTROPY

Traditional statistical mechanics defines the "Boltzmann" entropy of a classical system of distinguishable particles in terms of the logarithm of a volume in phase space.² For an ideal gas this definition leads to the equation^{3,4}

$$S_{\text{trad,dist}} = kN \left[\ln V + \frac{3}{2} \ln \frac{E}{N} + X \right], \quad (1)$$

where k is Boltzmann's constant, N is the number of particles, V is the volume, E is the energy, and X is a universal constant.

As pointed out by Gibbs, Eq. (1) for the entropy is not extensive⁵ and hence he argued that Eq. (1) is incorrect. However, many textbooks take the view that the entropy of a system of distinguishable particles is not extensive.^{3,4} It does not seem to be generally known that this lack leads to strange predictions for the properties of colloids.

III. THE PROBLEM WITH THE TRADITIONAL DEFINITION OF ENTROPY

Consider homogenized milk as an example. Homogenization reduces the size of the suspended globules of butterfat in milk to about 0.5 microns, which is small enough to prevent them from rapidly separating out. There are roughly 10^{12} globules of butterfat per liter of whole milk, which is sufficient for the application of statistical mechanics. To simplify the discussion without losing the essential features, we approximate milk as a dilute system, neglect the interactions between the globules of fat, and treat them as an ideal gas. The traditional form for the entropy would then be given by Eq. (1).

Imagine a container with a removable wall dividing it into two subvolumes, V_1 and V_2 , such that $V_1 + V_2 = V$. Let the total volume V be 1 liter. Take a liter of milk and fill both subvolumes. Let N_1 and N_2 be the numbers of fat globules in each subvolume, with $N = N_1 + N_2$. Clearly, the densities of fat globules are the same in each subvolume, so that $N_1/V_1 = N_2/V_2 = N/V$. Now remove the wall. The change in the total traditional entropy

$$S_{\text{trad,dist,total}} = S_{\text{trad,dist,1}} + S_{\text{trad,dist,2}} \quad (2)$$

is

$$\Delta S_{\text{trad,dist,total}} = kN \ln V - kN_1 \ln V_1 - kN_2 \ln V_2 \quad (3a)$$

$$= kN_1 \ln V + kN_2 \ln V - kN_1 \ln V_1 - kN_2 \ln V_2 \quad (3b)$$

$$= kN_1 \ln \frac{V}{V_1} + kN_2 \ln \frac{V}{V_2}. \quad (3c)$$

The change of the entropy in Eq. (3c) is clearly greater than zero, so the entropy has increased, suggesting an irreversible process. However, if we now replace the wall, the system will return to the original macroscopic state and the original entropy. Because the final macroscopic state is the same as the initial one, replacing the wall requires a *decrease* in entropy in direct violation of the second law of thermodynamics.

The contradictions are just as unfortunate for other experiments we could envision. Suppose that $V_1 > V_2$, and we have poured whole milk into V_1 and skim milk (no suspended fat particles) into V_2 . Initially, $N_1 = N$ and $N_2 = 0$. The initial traditional entropy would be

$$S_{\text{trad,dist,initial}} = kN \left[\ln V_1 + \frac{3}{2} \ln \frac{E}{N} + X \right]. \quad (4)$$

Make a hole in the wall to let the whole milk and the skim milk mix for a short time, but close the hole before the milk in the two subvolumes has completely mixed. The densities in the subvolumes now satisfy the inequalities

$$\frac{N_1}{V_1} > \frac{N}{V} > \frac{N_2}{V_2}. \quad (5)$$

The change in the traditional entropy for this experiment is

$$\Delta S_{\text{trad,dist,total}} = kN_1 \ln V_1 + kN_2 \ln V_2 - kN \ln V_1 \quad (6a)$$

$$= -kN_2 \ln V_1 + kN_2 \ln V_2 \quad (6b)$$

$$= -kN_2 \ln \frac{V_1}{V_2}. \quad (6c)$$

Because I specified that $V_1 > V_2$, Eq. (6c) implies that the traditional entropy will decrease, $\Delta S_{\text{trad,dist,total}} < 0$, even though the process is clearly irreversible and the entropy should increase.

IV. HOW DID BOLTZMANN REALLY DEFINE THE ENTROPY?

We can see from these examples as well as others given in Ref. 6 that the traditional expression for the entropy is untenable. Oddly enough, Boltzmann would not have encountered these problems, because he would not have used Eq. (1). He wrote the entropy (in modern notation) as⁷

$$S_{\text{dist}} = kN \left[\ln \frac{V}{N} + \frac{3}{2} \ln \frac{E}{N} + X + 1 \right]. \quad (7)$$

If we use Eq. (7), the entropy remains constant in the first experiment when the wall between the two subvolumes of milk is either removed or reinserted, as is appropriate for a reversible process. For the second experiment, it is easy to show that $\Delta S_{\text{dist,total}}$ is always positive; the entropy increases as it must for an irreversible process.

How did we arrive at a situation in which traditional sources give an entropy that is both incorrect and different from what Boltzmann had written? I believe that the development of quantum mechanics led people to interpret Boltzmann's classical papers differently than he had intended. The quantum concept of indistinguishable particles presented such a wonderful excuse to divide by $N!$ that it was assumed that Boltzmann would have used it. I disagree with this assumption and the disagreement becomes important for colloids.

This issue is also closely connected with the incorrect but widespread belief that Boltzmann defined the entropy as the logarithm of a volume in phase space. In his 1877 paper Boltzmann gave his interpretation of the second law and the entropy.

“Der Anfangszustand wird in den meisten Fällen ein sehr unwahrscheinlicher sein, von ihm wird das System immer wahrscheinlicheren Zuständen zueilen, bis es endlich den wahrscheinlichsten, d.h. den des Wärmegleichgewichtes, erreicht hat. Wenden wir dies auf den zweiten Hauptsatz an, so

*können wir diejenige Größe, welche man gewöhnlich als die Entropie zu bezeichnen pflegt, mit der Wahrscheinlichkeit des betreffenden Zustandes identifizieren.”*⁸

The initial state will usually be a very improbable one. The system will run from this state to ever more probable states until it reaches the most probable one, which is the state of thermal equilibrium. If we apply this to the second law, we can identify the quantity that is usually called the entropy with the probability of that state.⁹

Boltzmann also stressed the importance of analyzing a composite system in defining the entropy.¹⁰ Following the previous quote, Boltzmann continued:

*“Denken wir uns ein System von Körpern, welche für sich isoliert und nicht mit anderen Körpern in Wechselwirkung sind, z.B. einen Körper von höherer und einen von niedererer Temperatur oder Spannung, und einen sogenannten Zwischenkörper, welcher die Wärmeübertragung zwischen beides vermittelt, oder um ein anderes Beispiel zu wählen, ein Gefäß mit absolut glatten und starren Wänden, dessen eine Hälfte mit Luft von geringerer Temperatur oder Spannung, dessen anderer Hälfte mit Luft von höherer Temperatur oder Spannung erfüllt ist. Das System von Körpern, welches wir uns gedacht haben, habe zu Anfang der Zeit irgend einen Zustand; durch die Wechselwirkung der Körper verändert sich dieser Zustand; gemäß dem zweiten Hauptsatze muß diese Veränderung immer so geschehen, daß die gesamte Entropie aller Körper zunimmt; nach unserer gegenwärtigen Interpretation heißt dies nichts anderes, als die Wahrscheinlichkeit des Gesamtzustandes aller dieser Körper wird immer größer; das System von Körpern geht stets von einem unwahrscheinlicheren zu einem wahrscheinlicheren Zustande über.”*⁸

Consider a composite system that is isolated and does not interact with other objects. For example, consider an object at higher temperature or pressure and another at a lower temperature or pressure, with another object that transmits heat between them. Or, to take another example, consider a container with absolutely smooth and rigid walls. One half is filled with air at a lower temperature or pressure and the other half is filled with air at a higher temperature or pressure. This composite system begins in an arbitrary state and by its interactions changes its state. According to the second law, these changes must always increase the total entropy of the entire system. According to our present interpretation, this means nothing other than that the probability of the complete state of all

the objects constantly increases. The composite system goes steadily from a more improbable state to a more probable state.⁹

Boltzmann identified the entropy with the probability of the macroscopic state of a composite system. Although he did not say here that the entropy is the logarithm of the probability, this association appears later in the same paper when he defined a quantity that he calls the “*Permutabilitätsmaß*” (measure of permutations).⁸ Because Boltzmann calculated probabilities by discretizing the relevant variables and counting permutations, this quantity is essentially the logarithm of the probability of the macroscopic state. He identifies the *Permutabilitätsmaß* with the entropy for a special case (to within additive and multiplicative constants).

I believe that this identification is the point in Boltzmann’s argument that has led people to believe that he defined entropy in terms of the logarithm of a volume in phase space. The *Permutabilitätsmaß* would be proportional to such a quantity, and using it as a definition leads to the expression for entropy given in Eq. (1). This identification left systems of distinguishable particles without the factor of $1/N!$ that Boltzmann (and Gibbs⁵) had included.

V. RESOLUTION OF THE PROBLEM

Once it is recognized that the entropy should be defined as the logarithm of a probability, it is easy to see why earlier workers had difficulty in explaining the factor of $1/N!$. It looks as if the number of states has been *overcounted*, when it has really been *undercounted*. Expressing the probability in Boltzmann’s language of permutations, when a system of distinguishable particles is allowed to exchange particles with the rest of the world, we must include the permutations of all possible combinations of particles that might enter or leave the system. This number is enormous, so the traditional expression seriously undercounts the permutations. Fortunately, correct counting leads to physically sensible results, including the factor of $1/N!$ for both distinguishable and indistinguishable classical particles.⁶

We can understand the basic idea by neglecting the energy dependence and considering the probability distribution of distinguishable ideal gas particles between two subvolumes. We have

$$P(N_1, N_2) = \frac{N!}{N_1! N_2!} \left(\frac{V_1}{V}\right)^{N_1} \left(\frac{V_2}{V}\right)^{N_2}, \quad (8)$$

where $P(N_1, N_2)$ is the probability of N_1 particles in volume V_1 and N_2 particles in volume V_2 . In Eq. (8), $N_1 + N_2 = N$ is the fixed total number of particles and $V_1 + V_2 = V$ is the total volume. Following Boltzmann, we define the entropy of this composite system by

$$S_{\text{dist}}(N_1, V_1, N_2, V_2) = k \ln P(N_1, N_2) + C, \quad (9)$$

where the constant C may depend on the constants N and V but not on the variables N_1 , N_2 , V_1 , and V_2 .

We want the composite entropy to equal the sum of the entropies of each subsystem in the thermodynamic limit. Thus, the constant C should cancel out the N and V dependence of $k \ln P(N_1, N_2)$ so that the composite entropy has no explicit dependence on N and V and depends only on the

subvolume variables. Because the N and V dependence of $k \ln P(N_1, N_2)$ is given by $-k \ln(V^N/N!)$, we let $C = k \ln(V^N/N!)$. In this way we find

$$S_{\text{dist}}(N_1, V_1, N_2, V_2) = k \ln P(N_1, N_2) + k \ln \frac{V^N}{N!} = k \ln \frac{V_1^{N_1}}{N_1!} + k \ln \frac{V_2^{N_2}}{N_2!}. \quad (10)$$

We introduce the entropy function $S(N, V) = k \ln(V^N/N!)$ and write (10) as

$$S_{\text{dist}}(N_1, V_1, N_2, V_2) = S(N_1, V_1) + S(N_2, V_2). \quad (11)$$

The composite entropy in Eq. (11) is the sum of the entropies of the subsystems.

For a large system in equilibrium, $N_1 = NV_1/V$ and $N_2 = NV_2/V$, so that $\ln P$ is of order $\ln N$, but the constant C is of order $N \ln N$, which is much larger. Thus, neglecting terms of order $1/N$, we find that in equilibrium

$$S_{\text{dist}}(N_1, V_1, N_2, V_2) = S(N_1, V_1) + S(N_2, V_2) \rightarrow S(N, V). \quad (12)$$

The functional form of the entropy for the composite system when the constraints are removed is the same as that for each subsystem. This physically desirable feature follows from the identification of the entropy with the logarithm of the probability in Eq. (9) plus a constant.

The general relation between a composite system of distinguishable classical particles and a corresponding system with the same interactions but indistinguishable classical particles turns out to be quite simple. Let the probability distribution for the number of indistinguishable particles in a general composite system be written as

$$P_{\text{indist}}(E_1, N_1, V_1, E_2, N_2, V_2) = \frac{\Omega_1(E_1, N_1, V_1) \Omega_2(E_2, N_2, V_2)}{\Omega(E, N, V)}, \quad (13)$$

where I have used the fact that the probability can be expressed by two factors that depend on the variables in each subsystem (denoted as Ω_1 and Ω_2) and a normalization factor that depends only on the total variables (denoted as Ω).⁶ The entropy is then given by

$$S_{\text{indist}}(E_1, N_1, V_1, E_2, N_2, V_2) = k \ln P_{\text{indist}}(E_1, N_1, V_1, E_2, N_2, V_2) + k \ln \Omega(E, N, V) \quad (14a)$$

$$= k \ln \Omega_1(E_1, N_1, V_1) + k \ln \Omega_2(E_2, N_2, V_2), \quad (14b)$$

where I have added a constant, $k \ln \Omega(E, N, V)$ that depends only on E , N , and V . If we make a similar identification as before, we have

$$S_{\text{indist}}(E_1, N_1, V_1, E_2, N_2, V_2) = k \ln P_{\text{indist}}(E_1, N_1, V_1, E_2, N_2, V_2) + S(E, N, V) \quad (15a)$$

$$= S_1(E_1, N_1, V_1) + S_1(E_2, N_2, V_2), \quad (15b)$$

where $S_j(E_j, N_j, V_j) \equiv k \ln \Omega_j(E_j, N_j, V_j)$.

The number of states for a system of N distinguishable

particles is equal to $N!$ times the number of states in an equivalent system of N indistinguishable particles. Thus, to obtain the probability distribution for distinguishable particles, we must modify Eq. (13) by letting $\Omega_j(N_j)$

$\rightarrow \Omega_j(N_j)N_j!$ and $\Omega(N) \rightarrow \Omega(N)N!$. In addition we must multiply the expression in Eq. (13) by $N!/N_1!N_2!$ to account for permutations of the distinguishable particles between the subsystems. This reasoning gives

$$P_{\text{dist}}(E_1, N_1, V, E_2, N_2, V_2) = \frac{\Omega_1(E_1, N_1, V_1)N_1!\Omega_2(E_2, N_2, V_2)N_2!}{\Omega(E, N, V)N!} \frac{N!}{N_1!N_2!} \quad (16a)$$

$$= \frac{\Omega_1(E_1, N_1, V_1)\Omega_2(E_2, N_2, V_2)}{\Omega(E, N, V)} \quad (16b)$$

$$= P_{\text{indist}}(E_1, N_1, V_1, E_2, N_2, V_2). \quad (16c)$$

If we take the logarithm of Eq. (16c), we see that $S_{\text{dist}}(E, N, V) = S_{\text{indist}}(E, N, V)$ for classical systems.

VI. CONCLUSIONS

Boltzmann's genius was seen early in his career. Even those of his contemporaries who argued against his ideas recognized his brilliance. Unfortunately, his ideas were not always completely understood and therefore not always completely appreciated.

In the preface to his book on Boltzmann's life and work, Cercignani commented that:

"It is remarkable that, with a few exceptions, Boltzmann's scientific papers have not been translated into English, whereas this task has been accomplished for other scientists of equal or lesser importance. Because of this, much of Boltzmann's work is known through somebody else's presentation, not always faithful."¹¹

Secondary sources have dominated Boltzmann's legacy to such an extent that even the famous equation inscribed on his tombstone, $S = k \log W$, is written in a form due to Max Planck.^{12,13} Planck did correctly choose the notation W to indicate "Wahrscheinlichkeit" (probability).

I was not aware of Boltzmann's 1877 identification of the entropy with the probability of macroscopic states of a composite system until quite recently. In a previous paper,⁶ I followed secondary sources in making the error of attributing the shortcomings of Eq. (1) to Boltzmann. Although the identification of the entropy with the logarithm of a volume in phase space did originate with Boltzmann, it was only a special case. Boltzmann's fundamental definition of the entropy in his 1877 paper has none of the shortcomings resulting from applying an equation for a special case beyond its range of validity.

By returning to Boltzmann's fundamental definition, I believe that we gain a deeper insight into the meaning of the entropy. Gibbs' paradox disappears. The equations for both distinguishable and indistinguishable particles take on clearer interpretations. Finally, we are led to a consistent statistical mechanics and thermodynamics for the description of colloids, which have too long been neglected.

As we approach the 100th anniversary of Boltzmann's sad death by his own hand on 5 September 1906, it is fitting to recognize that his insights into the foundations of statistical physics even exceeded those for which he is already justly famous.

¹D. H. Everett, *Basic Principles of Colloid Science* (Royal Society of Chemistry, London, 1988).

²See, for example, L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press, New York, 1980), pp. 24–25.

³F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965), pp. 243–245.

⁴R. K. Pathria, *Statistical Mechanics* (Butterworth-Heinemann, Boston, 1996), pp. 22–26. See especially Pathria's Eq. (2a) on page 23, which agrees with my Eq. (1) if the temperature is replaced by the energy as a variable.

⁵J. W. Gibbs, *Elementary Principles of Statistical Mechanics* (Yale University Press, New Haven, 1902). (Reprinted by Dover, New York, 1960), pp. 206–207.

⁶R. H. Swendsen, "Statistical mechanics of classical systems of distinguishable particles," *J. Stat. Phys.* **107**, 1143–1165 (2002).

⁷L. Boltzmann, *Vorlesungen über Gastheorie* [J. A. Barth, Leipzig, Germany, 1896 (Part I) and 1898 (Part II)], translated into English by Stephen G. Brush in L. Boltzmann, *Lectures on Gas Theory* (Dover, New York, 1995), p. 72.

⁸L. 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. The quotations are to be found on pp. 165 and 166.

⁹Because I was not able to find English translations of Boltzmann's papers, the translations in this paper are my own. I have tried to be faithful to Boltzmann's meaning, but not to his sentence structure.

¹⁰N. G. van Kampen, "The Gibbs paradox," in *Essays in Theoretical Physics*, edited by W. E. Parry (Pergamon, Oxford, 1984). The importance of considering composite systems was also pointed out by van Kampen in this article.

¹¹C. Cercignani, *Ludwig Boltzmann: The Man Who Trusted Atoms* (Oxford University Press, New York, 1998), p. xi.

¹²M. Planck, "Über das Gesetz der Energieverteilung im Normalspektrum," *Drudes Annalen*, 553–562 (1901), reprinted in *Ostwalds Klassiker der exakten Wissenschaften, Band 206*, "Die Ableitung der Strahlungsgesetze," pp. 65–74. The equation is given with an arbitrary additive constant on p. 68 of the reprinted text.

¹³M. 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.