Entropy, Its Language, and Interpretation

Harvey S. Leff

Abstract The language of entropy is examined for consistency with its mathematics and physics, and for its efficacy as a guide to what entropy means. Do common descriptors such as *disorder*, *missing information*, and *multiplicity* help or hinder understanding? Can the language of entropy be helpful in cases where entropy is not well defined? We argue in favor of the descriptor *spreading*, which entails space, time, and energy in a fundamental way. This includes *spreading* of energy spatially during processes and *temporal spreading* over accessible microstates states in thermodynamic equilibrium. Various examples illustrate the value of the spreading metaphor. To provide further support for this metaphor's utility, it is shown how a set of reasonable spreading properties can be used to *derive* the entropy function. A main conclusion is that it is appropriate to view entropy's symbol S as shorthand for *spreading*.

1 Introduction

Thermodynamics and statistical mechanics were developed to describe macroscopic matter. They differ from mechanics, which describes point particles and rigid bodies, in that they account for internal energy storage modes. Accordingly, a key function of thermodynamics is internal energy, namely, the average *total* energy of a system, including all storage modes. An equally important, but less easily digested, function is entropy. Perhaps because it can be defined in diverse ways, and its behavior in thermodynamic systems can be subtle, entropy's meaning and usefulness have come into question [1]. A commonly used, but inadequate, language surrounding entropy contributes to this. Despite its intimate connection with energy, entropy has been described as a measure of disorder, multiplicity, missing information, freedom, mixed-

H.S. Leff (✉)
Department of Physics, California State Polytechnic University, 3801 W. Temple Ave., Pomona, CA 91768, USA
e-mail: hsleff@csupomona.edu

up-ness, and the like—none of which involves energy explicitly.[1] The purpose of the present article is to discuss how entropy can be related to energy in a qualitative, interpretive way, offering the possibility of an improved understanding and appreciation of entropy's generality and value.

1.1 The Disorder Metaphor

The most common metaphor for entropy relates it to *disorder*. Usage goes back at least to Boltzmann [2], who wrote, "...one must assume that an enormously complicated mechanical system represents a good picture of the world, and that all or at least most of the parts of it surrounding us are initially in a very ordered—therefore very improbable—state ... whenever two or more small parts of it come into interaction with each other, the system formed by these parts is also initially in an ordered state, and when left to itself it rapidly proceeds to the disordered most probable state."

Unfortunately, the term disorder is deficient. First and foremost, there is no *general* definition of it in the context of thermodynamics. Dictionary definitions typically are: *Lack of order or regular arrangement; confusion; (in medicine) a disturbance of normal functioning.* "Lack of regular arrangement" has a spatial connotation, and one can indeed conceive of *spatial* disorder; i.e., where the positions of particles lack regularity; and *spin* disorder, which refers to the degree of non-regularity of up-down spin orientations. The term *confusion* can be related to the disorder that some associate with higher temperatures, often envisaged in terms of increased *thermal agitation*. The definition of disorder seems to be variable, depending on the situation, which makes the term vague and confusing. Burgers [3] observed that disorder "leads to anthropomorphic forms of reasoning which are often misleading." Yet physicists still tend to gravitate toward use of the disorder metaphor.

Another difficulty is that one can be misled by focusing attention on one obvious type of disorder to the exclusion of important others. For example, this is tempting in the discussion of spontaneous crystal formation in an isolated, metastable supersaturated solution. Here a dissolved solid's concentration exceeds its maximum equilibrium value and a slight perturbation can trigger crystal formation. From a spatial viewpoint, there is more *order* after crystallization, yet the system's entropy (including liquid, solid crystals, and container walls) must increase. The tempting misleading interpretation that entropy decreases, based on increased spatial order is exacerbated by the fact that it is possible that temperature has also decreased. To be sure, the physical phenomena here are not transparent, but the ease of misinterpretation using the disorder metaphor is noteworthy.

The behavior of some binary liquid crystals clearly illustrate weaknesses with the disorder metaphor [4]. At one temperature, such a system can be in a *nematic* phase, where the rod-like molecules tend to align, but positions show no particular spatial order. For lower temperatures, the system can enter the *smectic* phase, where molecules lie in well defined planes, establishing a spatial layering. The increased spatial order

---

[1]Some entail energy implicitly. For example, multiplicity refers to the number of accessible *energy* states consistent with specified conditions.

(decreased disorder) does indeed accompany reduced entropy. However, as temperature is lowered further, the system re-enters a *nematic* phase with *increased* spatial disorder, but *lower* entropy. This shows again that the disorder metaphor can easily mislead. Generally, an undue focus on *configurational* entropy (including that associated with the regularity of positions or alignment of rodlike or polarized molecules), without giving proper attention to entropy contributions from linear and angular momentum effects, is a dangerous practice.

Karl Darrow [5] examined examples for which entropy could be associated with disorder, and observed that disorder is not always useful, as indicated by a free expansion of a gas in a constant-temperature environment. Entropy increases, but one cannot unambiguously associate more disorder with the final state. He wrote, "Would anyone spontaneously say that a gas in a two-liter bottle is more disorderly than the same gas at the same temperature in a one-liter bottle?" In another discussion of the free expansion, Bohren and Aldrich [6] write that increased volume only means increased disorder if "we *define* this to be so. But if we do, we have defeated our purpose of supposedly explaining entropy as disorder. That is, we have defined disorder by means of entropy, not the other way around."

Despite Darrow's displeasure with some aspects of the disorder metaphor, he used it subsequently in its variable-definition form and applied it to the zero-temperature limit: "Zero entropy corresponds to perfect order ... if two or more types of disorder coexist each makes to the total entropy a contribution of its own, which vanishes when it vanishes." This implies that as entropy approaches zero, all possible forms of disorder disappear. Although this circumvents *defining* disorder for the zero-temperature limit, it does not provide much help with finite temperature situations, where definitions of various types of disorder would be needed.

The term disorder has been criticized by a variety of others writers, including Dingle [7], who described the view that entropy is disorder as "a most inessential visualization which has probably done much more harm than good." Wright [8] examined various examples of real phenomena and concluded, "...there is no clear correlation between propositions about entropy and anything *intuitively obvious* about disorder." Lambert [9], critically assessed usage of the disorder metaphor in chemistry textbooks, and has successfully convinced authors of 15 textbooks to remove statements relating entropy with *disorder* from new editions of their books.

Upon the death of J. Willard Gibbs, a list of subjects intended for supplementary chapters to Gibbs' *Equilibrium of Heterogeneous Substances* was discovered [10]. One of the topics was "On entropy as mixed-up-ness." Unfortunately, Gibbs did not live to bring this to fruition and it is not known just what Gibbs had in mind. Mixed-up-ness sounds a lot like disorder and Gibbs, who had considerable mathematical skills, might have been able to solidify its meaning. As it stands, even if one places value in the disorder metaphor, its current use does not entail specific reference to energy, the centerpiece of thermodynamics. In this sense, it cannot get to the heart of the physics.

1.2 Missing Information, Multiplicity, Optiony, Freedom, Unavailability

The metaphor of *missing information* for entropy is much preferred over disorder. Edwin T. Jaynes [11], used information theory to develop the full mathematical

framework of equilibrium statistical mechanics. The missing information metaphor is well defined and can be quite useful, especially in understanding that descriptions of macroscopic matter necessarily discard enormous amounts of information about system details, working ultimately with a small number of macroscopic variables such as pressure, volume, and temperature. It does not, however, use space, time, and energy in a qualitatively useful way, and thus cannot replace the spreading metaphor that is proposed herein as an interpretive tool.

Related terms are *multiplicity* [12] or equivalently, *optiony* [13]. These terms refer to the number of microstates that correspond to a given macrostate. With these choices, entropy is defined in terms of Boltzmann's famous $S = k_B \ln(\textit{multiplicity}) = k_B \ln(\textit{optiony})$, where k_B is Boltzmann's constant (the Boltzmann form is discussed further in Sect. 1.3). The Second Law can then be stated: *If an isolated macroscopic system is permitted to change, it will evolve to the macrostate of largest multiplicity (or optiony) and will remain in that macrostate*.

The metaphor *freedom* was proposed by Styer [4], who wrote, "... the advantage of the 'entropy as freedom' analogy is that it focuses attention on the variety of microstates corresponding to a macrostate whereas the 'entropy as disorder' analogy invites focus on a single microstate." Styer also warns of deficiencies in the term *freedom*, and suggests that one use *both* freedom and disorder together to better see the meaning of entropy. The freedom metaphor was proposed independently by Brissaud [14]. Freedom can be related to multiplicity, optiony and missing information, and has some attractive features.

Finally, there exists a common metaphor that entropy is a measure of the *unavailability* of energy that can be converted to work in some processes. Because the energy alluded to is macroscopic energy in this engineering-oriented definition, it cannot help us understand why, for example, 2 kg of copper has twice the entropy of 1 kg of copper.

Although the above terms can all be helpful, they do not convey the notions that thermodynamic processes entail energy spreading and that thermodynamic equilibrium is a dynamic equilibrium at a microscopic level.

1.3 The 'Spreading' Metaphor

The metaphor of *spreading* is based explicitly upon space, time, and energy. Space is intimately involved because energy tends to *spread* spatially to the extent possible. For example when hot and cold objects are put in thermal contact, energy spreads *equitably* (discussed in Sect. 3) between them. And when an object is in thermal equilibrium at constant temperature, its quantum state varies temporally as the system's state point *spreads* over accessible quantum states that are consistent with the thermodynamic state. Spatial spreading is a helpful interpretive tool for *processes* and *equilibrium states*, and temporal spreading is particularly useful for interpreting entropy in a particular thermodynamic state.

Clausius came close to using the concept of spreading even before he published his path-breaking 1865 paper that introduced entropy. In 1862 he proposed a function that he called *disgregation* [15]. Clausius pictured molecules as being in constant motion

and viewed this *heat*[2] as tending to "loosen the connection between the molecules, and so to increase their mean distances from one another." Clausius went further in his 1865 introduction of entropy, writing

$$S - S_o = \int \frac{dQ}{T} = \int \frac{dH}{T} + \int dZ. \tag{1}$$

This represents entropy relative to its value (S_o) in a reference state as the sum of two terms. In the first, dH is the change in what Clausius called *heat content*, which we now call the average kinetic energy as calculated in the canonical ensemble of classical statistical mechanics. In the second term, dZ is the change in disgregation, which we now understand [16] comes from the position integrals of the canonical partition function in classical statistical mechanics. This is consistent with disgregation being related to what is referred to herein as *spatial spreading*.

Denbigh [17] used the spreading idea to describe irreversibility, citing three forms that display *divergence toward the future*: (i) a branching towards a greater number of distinct kinds of entities; (ii) a divergence from each other of particle trajectories, or of sections of wave fronts; (iii) a spreading over an increased number of states of the same entities. These statements entail space and time and although they do not refer specifically to *energy*, they can easily be interpreted in terms of it. Earlier, Denbigh wrote [18], "As soon as it is accepted that matter consists of small particles which are in motion it becomes evident that every large-scale natural process is essentially a process of mixing, if this term is given a rather wide meaning. In many instances the spontaneous mixing tendency is simply the intermingling of the constituent particles, as in inter-diffusion of gases, liquids and solids.... Similarly, the irreversible expansion of a gas may be regarded as a process in which the molecules become more completely mixed over the available space.... In other instances it is not so much a question of a mixing of the particles in space as of a mixing or sharing of their total energy." To this it should be added that when particles move and mix, they carry with them their translational kinetic and stored (e.g., rotational or vibrational) energies; i.e., they spread their energies and exchange energy with other particles.

While multiplicity, missing information, and the like entail the number of *possible* states, the spreading metaphor entails a picture of *dynamic* equilibrium in terms of continual shifts from one microstate to another. The difference can be viewed as use of a noun (multiplicity, information, ...) *vs.* use of a verb (spreading). The spreading route envisages an *active* system, where macroscopically invisible energy exchanges take place—even in equilibrium—while the alternative noun descriptors picture what is possible, rather than what happens spatially and temporally. The mathematics is the same for both, but the interpretations differ significantly.

There seems to be a tendency for some people to link uncertainty with disorder. For example, after a free expansion, the average volume per particle is larger, and we are less certain about where a particle is at any instant. In this sense, missing information has increased. Each molecule carries its translational and internally stored energies to a larger spatial region, but if the gas temperature is unchanged, or has

---

[2]This language is obsolete. The term *heat* is now reserved for energy transfers induced by temperature gradients.

decreased, in what sense has the gas become more disordered? Evidently some relate their own uncertainties about a particles' whereabouts with the degree of disorder in the gas itself. The poorly defined term *disorder* has anthropomorphic underpinnings that seem to make it acceptable for some to bend its meaning with its use.

The spreading metaphor is powerful, and offers a physically motivated, transparent alternative to the metaphors discussed above. Styer [4] observed that the most important failing of "...the analogy of entropy and disorder invites us to think about a single configuration rather than a class of configurations." The spreading metaphor addresses this deficiency by focusing on temporal shifts between configurations. In 1996, the spreading metaphor was used to introduce and motivate the development of thermodynamics pedagogically [19]. In 2002, Lambert [9, 20] argued in favor of energy-based language to replace the inadequate metaphor of disorder in chemistry textbooks. The present paper is intended to extend the discussion of the spreading metaphor for a wider audience of research scientists.

2 Spreading in Equilibrium Thermodynamic States

2.1 Following Boltzmann's Lead

Microscopic connections between entropy and spreading can be seen best via the Boltzmann entropy expression

$$S = k_B \ln W(E, V, N), \tag{2}$$

where k_B is Boltzmann's constant and $W(E, V, N)$ is the number of microstates that are consistent with total system energy E, volume V, and particle number N. Classically, $W(E, V, N)$ is an integral over the $6N$-dimensional phase space [21],

$$W(E, V, N) = \frac{1}{N!} \int \left(\frac{d^3q \, d^3p}{h^3} \right)^N \delta(E - H_N), \tag{3}$$

where H_N is the system's Hamiltonian. The delta function $\delta(E - H_N)$ restricts nonzero contributions to the integrals to a $6N - 1$ dimensional constant-energy surface. These equations are the basis of the microcanonical ensemble formalism of statistical mechanics, from which the more widely used canonical, and grand canonical ensemble formalisms are derived. It is natural to use the fundamentally important (2) and (3) to introduce the concept of *spreading*.

The space and momentum integrals run over all possible spatial positions and momenta for each particle. By varying initial conditions, *any* point in the phase space volume $W(E, V, N)$ is accessible.[3] The functions $W(E, V, N)$ and $S(E, V, N)$ therefore increase with E and V, holding the remaining two variables fixed, because an increase in either of these variables increases the accessible phase space volume.[4]

---

[3]It is implicitly assumed that accessible phase space regions with equal phase space volumes are equally likely and that equal times are spent in them.

[4]Changes with N at fixed E and V are more subtle, and we *cannot* conclude that $S(E, V, N)$ always increases with N.

A single system's phase point traverses the phase space in time as particles move and exchange energy. This traversal provides a graphic image of a system's particles spreading and exchanging energy through space, as time progresses. Energy, space, and time are all explicitly involved. In this sense, entropy can be thought of as a *spreading* function.

In real systems, the total energy is never exactly constant and energy exchanges with the environment, even if weak, cause continual changes in the system's microstate. To address this case, the function $W(E, V, N)$ in (2) can be replaced by $W(E, \Delta E, V, N)$, the phase space volume of the set of points $\{q, p\}$ such that $E - \frac{1}{2}\Delta E <$ system energy $< E + \frac{1}{2}\Delta E$. This is the phase space volume of points in an energy shell rather than on an energy surface. Here energy exchanges with the environment cause continual changes in the system's phase point, in addition to the constant energy flow envisaged originally.

For a quantum system, $W(E, V, N)$ is the number of quantum states with total energy E; i.e., the degeneracy of E. In principle, the system can be in any of these states, or more generally a superposition of them. For an ensemble, if each system's state could be measured, one would find many of the possible states. This can be interpreted as a kind of spreading over the members of the ensemble. For a real system that interacts with its environment, as in the classical case, the number of states, $W(E, \Delta E, V, N)$, in the energy interval $(E - \frac{1}{2}\Delta E, E + \frac{1}{2}\Delta E)$ can be used. The interactions induce continual jumps from one quantum state to another, a veritable 'dance' over accessible states. This is *temporal spreading*, and this view provides an interpretation for a number of familiar properties of entropy.

For example, under given temperature and pressure conditions, the entropy per particle of monatomic ideal gases tend to increase with atomic mass [19, 22]. This can be understood because the energy levels of an ideal gas are inversely proportional to atomic mass and thus are closer together for higher mass atoms. Thus, in any energy interval ΔE, there will be more states for gases with higher atomic mass, and more states over which the dance over accessible states can occur. This increased *temporal* spreading over states corresponds to higher entropy.

Another example is a comparison of the entropy per particle $S(T, p, N)$ of N-particle monatomic and polyatomic gases at temperature T and pressure p. It is found that $S_{\text{mono}}(T, p, N) \leq S_{\text{poly}}(T, p, N)$. This a consequence of there being more states in an energy interval ΔE in the neighborhood of the average system energy because of the existence of rotations and vibrations in polyatomic molecules. Furthermore, substances comprised of polyatomic molecules with more atoms/molecule tend to have higher entropy values [22] for the same basic reason. More degrees of freedom lead to more states in a given energy interval. This implies a greater degree of temporal spreading over microstates and higher entropy.

2.2 Maxwell Momentum Distribution

Spreading can be viewed in various ways, one of which is through the Maxwell momentum distribution $F(p)$, as illustrated in Fig. 1. Here momenta are separated into bins, in the spirit of the energy cells used by Boltzmann to enable combinatoric counting arguments. The momentum distribution is shown rather than the more

Fig. 1 Probability density of
gas particles as a function of
momentum for three
temperatures
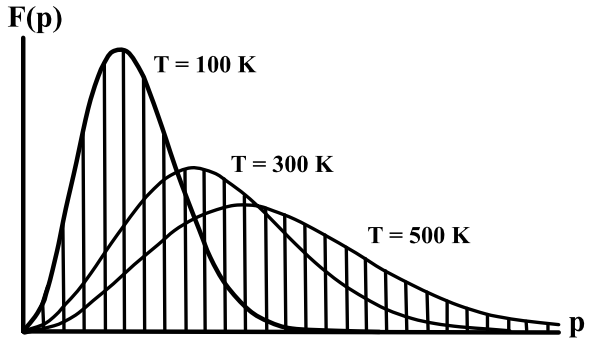
familiar speed distribution because momentum is a more informative variable, con-
taining both speed and mass information—and furthermore, it is the variable used,
along with position, in the phase space. Examination of the curves shows that, as ex-
pected, more higher-momentum bins become significantly occupied as temperature
increases. As gas molecules exchange energy with one another and with container
walls, they shift from bin to bin. This reflects a temporal spreading similar to that
described above in terms of the phase space trajectory for classical systems and for
a dance over discrete energy states for quantum systems. More temporal spreading
over bins at higher temperatures means higher entropy.

2.3 Effects of Interparticle Forces

The existence of interparticle forces can inhibit spreading and lower the value of
$W(E, V, N)$ and thus the corresponding entropy. This effect can be appreciated
by examining a one-dimensional classical harmonic oscillator with energy $E =
p^2/2m + kx^2/2$. For fixed energy E, the momentum is bounded by $|p| \leq (2mE)^{1/2}$
and the position is bounded by $|x| \leq (2E/k)^{1/2}$. The system's phase space trajectory
is an ellipse and its phase space *volume* is the ellipse's circumference. The semi-axes
have lengths $(2mE)^{1/2}$ and $(2E/k)^{1/2}$. Although an exact closed form expression for
the circumference does not exist, it is clear that in the limit of zero force constant k,
the circumference approaches infinite length and as k becomes large the circumfer-
ence approaches a minimum for fixed E. Here $W(E, V, N)$ is independent of V, and
$N = 1$, so the simple argument here along with (2) imply that S is a decreasing func-
tion of k for arbitrarily high and low values. Of course the functions W and S are
of questionable value here, but this example does illustrate how a force can inhibit
spreading.

In classical statistical mechanics, a gas with temperature T, volume V, and par-
ticle number N, has maximum entropy when there are zero interparticle forces; i.e.,
$S_{\text{ideal}}(T, V, N) \geq S_{\text{nonideal}}(T, V, N)$. To see how this comes about, suppose the sys-
tem has Hamiltonian $H = K + \Phi$, where K and Φ are the kinetic and potential
energies. In the canonical ensemble, the entropy for a system with Hamiltonian H is
$S_H = k_B \ln Q_H + \langle H \rangle / T$. Here $Q_H = Tr[\exp(-H/(k_B T))]$, and for classical sys-
tems, the trace is taken to be the integral operator $(h^3 N!)^{-1} \int (d^3 q \, d^3 p)^N$. For an ideal
gas at the same temperature and volume, the entropy is $S_K = k_B \ln Q_K + \langle K \rangle / T$.

The Gibbs–Bogoliubov inequality [23], which compares two systems with Hamiltonians A and B, is $-k_B T \ln Q_B \leq k_B T \ln Q_A + Q_A^{-1} Tr[(B-A) \exp(-A/(k_B T))]$. Choosing $A \equiv H$ and $B \equiv K$, the inequality implies $S_K - S_H \geq T^{-1}[\langle K \rangle_K - \langle K \rangle_H]$. In classical statistical mechanics, it is easy to see that $\langle K \rangle_K = \langle K \rangle_H$, showing that $S_K \geq S_H$, and thus, as stated above,

$$S_{\text{ideal}}(T, V, N) \geq S_{\text{nonideal}}(T, V, N). \tag{4}$$

Equation (4), derived by Baierlein [24] and generalized by Leff [25], reflects the fact that position-dependent forces reduce the degree of spatial spreading and the degree of temporal spreading as particles exchange energy. This occurs in the sense that correlations exist between the positions of interacting molecules, and these molecules are not as free to spread their energy as they are in an ideal gas. Equation (4) holds for charged particles in a magnetic field and for lattices of discrete and continuous spins [25].

Despite the fact that one can give a plausible argument for (4) to hold in the quantum domain,[5] it does *not* do so in general [26]. A counterexample has been exhibited for a two-dimensional model of low-density helium monolayers and for ideal quantum gases at sufficiently low temperatures. An interpretation is that this reflects the existence of *effective forces* associated with Bose–Einstein and Fermi–Dirac statistics that can offset the effects of *electromagnetic* intermolecular forces—and thus the degree of energy spreading. The delicacy of the competition between quantum mechanical and electromagnetic forces is evident from the result that (4) is violated in helium monolayers for low temperatures sufficient for quantum effects to exist, but is actually satisfied for *even lower* temperatures.

Additionally, writing $E = K + \phi$, the well known constant-volume heat capacity relation, $k_B T^2 C_v = \langle (E - \langle E \rangle)^2 \rangle$ implies

$$k_B T^2 C_v = [\langle K^2 \rangle - \langle K \rangle^2] + [\langle \Phi^2 \rangle - \langle \Phi \rangle^2] + 2[\langle K \Phi \rangle - \langle K \rangle \langle \Phi \rangle]. \tag{5}$$

For position dependent potential energies Φ, in classical statistical mechanics the last bracket vanishes. The first bracket is $k_B T^2 C_{v,\text{ideal}}$, and the middle bracket is non-negative. Thus we find the pleasing inequality

$$C_{v,\text{nonideal}} \geq C_{v,\text{ideal}}. \tag{6}$$

In essence, the existence of potential energy, in addition to kinetic energy, provides a second mode for energy storage, increasing the heat capacity. Using (6), constant-volume integration between specified initial and final temperatures, provides the constant-volume entropy inequality,

$$(\Delta S_{\text{nonideal}})_v \geq (\Delta S_{\text{ideal}})_v. \tag{7}$$

---

[5]When the ideal gas energy states are highly degenerate, one expects the addition of interparticle forces to split this degeneracy, and thereby *lower* the system entropy [24]; i.e., there will be fewer states and thus less spreading in a given energy interval. Perhaps high enough degeneracy is guaranteed only for sufficiently high energy states, which dominate only in the classical domain.

Here is an interpretation of (7) for heating. At higher temperatures, forces become less effective and spatial spreading in the nonideal gas entropy approaches that for the ideal gas. Given the initially reduced spatial spreading in the nonideal gas, this accounts for (7).

2.4 Photon Gas

The photon gas in an otherwise empty container whose walls are at constant temperature T is an interesting system [27]. If the container is a cylinder and one of the cylinder's end walls is a movable piston, one can envisage (in principle) starting out with zero photons, with the piston making contact with the opposite (fixed) end. If the piston is slowly moved away from the fixed end, photons pour into the container from the walls, establishing average photon number $N(T, V) = rVT^3$, with $r = 2.03 \times 10^7 \ \text{m}^{-3} \ \text{K}^{-3}$, at volume V. The photon gas has literally been formed by movement of the piston, providing space for the photons. The corresponding internal energy is $U(T, V) = bVT^4$, where $b = 7.56 \times 10^{-16} \ \text{J} \ \text{K}^{-4} \ \text{m}^{-3}$. The pressure of the photon gas at volume V is $p = \frac{1}{3}bT^4$ and the work done by the gas in the volume change from zero to V is $W = \frac{1}{3}bVT^4$. The first law of thermodynamics then implies isothermal heat transfer $Q = \frac{4}{3}bVT^4$, and thus the final photon gas entropy is

$$S(T, V) = Q/T = \frac{4}{3}bVT^3. \tag{8}$$

Note that $S(T, V) \propto N(T, V)$.

In the process of building the gas, photons and their energies *spread* throughout the container, clearly increasing the spatial spreading of energy. Once established, the dance over accessible states can be understood in terms of the continual absorption and emission of photons over time. If T is increased, the average number N of photons increases, which increases the amount of temporal energy spreading associated with absorption and emission of photons. Additionally, Fig. 2 shows that the entropy density—namely, the entropy per unit volume, per unit frequency of the photon gas—changes with temperature in a manner reminiscent of the Maxwell momentum distribution discussed in Sect. 2.2. An increasing number of bins hold significant numbers

Fig. 2 Entropy per unit volume, per unit frequency of a photon gas for three temperatures
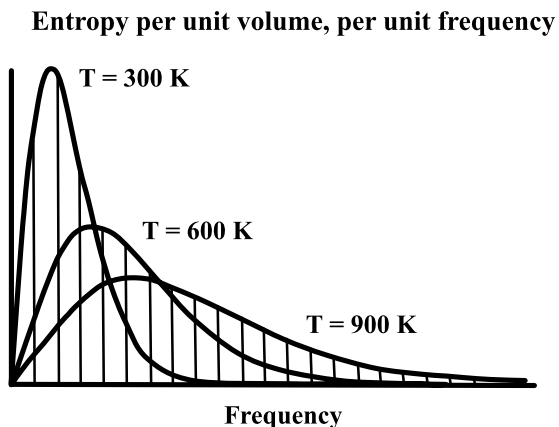
of photons at higher temperatures. This implies a greater degree of temporal spreading because of continual absorption and emission of photons and fluctuating numbers of photons and energy within more bins.

2.5 Thermal Wavelength and Spreading

When (2) is evaluated for a classical ideal gas, and the thermodynamics equation $(\partial S/\partial E)_V = 1/T$ is invoked, the result obtained in the large N limit is

$$S(T, V, N) = Nk_B[\ln(v/\lambda^3) + \text{constant}],\tag{9}$$

with $v \equiv V/N$, $\lambda = h/(2\pi mk_BT)^{1/2}$, the thermal wavelength.

The thermal wavelength can be viewed as the quantum extent of each particle and λ^3 as its quantum volume; i.e., an estimate of the volume of its wave packet. For $v > \lambda^3$, v/λ^3 is an indicator of the fraction of the container that is "occupied." Strictly speaking (9) holds only for $v \gg \lambda^3$; i.e., in the classical limit. Equation (9) predicts that S decreases as T is decreased with v fixed.

Commonly, the effects of temperature change are viewed in terms of momentum, but (9) suggests a very different view of the situation [19]. Suppose the total spatial volume V is divided into $M = V/\lambda^3$ cells, each the size of a wave packet. An atom can spread its energy to any of the M cells. The particles behave as independent entities, with negligible quantum interference if $M \gg N$; i.e., $\lambda^3 \ll v$. As T is lowered, M decreases and the number of accessible cells becomes smaller, reducing the degree of energy spreading in the sense that it becomes more likely, and eventually, necessary that some cells contain more than one particle. Increasing the temperature leads to the opposite conclusion, namely, the amount of energy spreading increases. All of this is consistent with the expected inequality, $(\partial S(T, V, N)/\partial T)_V > 0$.

Although (9) and the cell picture above hold for non-interfering particles only—i.e., for $\lambda^3 \ll v$, the thermal wavelength concept is valid for all temperatures. As T is lowered and λ increases, it becomes more likely that two or more particles will have overlapping wave packets; i.e., for $M < N$, some cells *must* be occupied by more than one particle. Let T_o denote the temperature at which $\lambda^3 = v$. At this temperature there are as many wave packet-sized cells as particles. As T is lowered below T_o, quantum interference builds, as illustrated symbolically in Fig. 3, and extreme quantum effects become possible. Such effects are expected for

$$T < T_o = \frac{h^2}{2\pi mk_B v^{2/3}}.\tag{10}$$

Bose–Einstein condensation is an extreme low-temperature behavior that has been verified experimentally [28]. Einstein's treatment of the ideal B-E gas shows a phase transition for *any* specific volume v at the critical temperature $T_{c,BE}$ such that $\lambda_c^3/v = 2.612$ [29]. This condition is consistent with expectations; i.e., $T_c < T_o$. Specifically,

$$T_{c,BE} = \frac{h^2}{2\pi mk_B(2.612)^{2/3}v^{2/3}} = 0.53T_o.\tag{11}$$

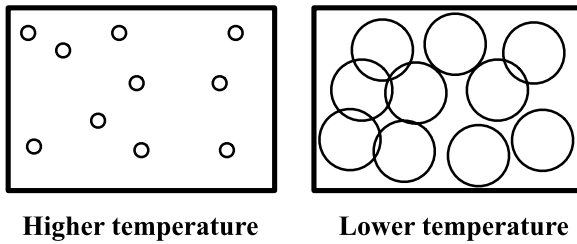**Higher temperature   Lower temperature**

Fig. 3 Symbolic representation of particle wave packets for higher temperatures, where $\lambda^3 \ll v$ and lower temperatures, where $\lambda^3 \approx v$. The rectangular boundaries represent spatial regions within the fluid (*not* the full container volume). The thermal wavelength λ does not become of macroscopic size until T is well below temperatures that have been reached in any laboratory

The main point here is that the growth of the thermal wavelength at low temperatures restricts spreading, albeit with some quantum subtleties.[6,7] The reduced spatial spreading signals reduced entropy and as T decreases toward absolute zero spatial spreading and entropy approach zero.

Although $\lambda \to \infty$ as $T \to 0$, in actual experiments λ does not become of macroscopic size until T is lower than temperatures that have been reached in any laboratory. For example, in the neighborhood of helium's lambda point, 2.7 K, where extreme quantum behavior is observed, $\lambda^3/v \approx 3$ (remarkably close to the condition for B-E condensation in an ideal gas) and $\lambda \approx 5 \times 10^{-10}$ m.

Nothing in our discussion of thermal wavelength restricts it to systems satisfying B-E statistics, and it should apply equally well to Fermi–Dirac statistics. In particular, (10) should still hold. For an ideal gas with spin 1/2 particles, which satisfies F-D statistics, the Fermi temperature is [29]

$$T_F = \frac{h^2}{2\pi m k_B v^{2/3}} \left(\frac{3\pi^{1/2}}{8} \right)^{2/3} = 0.76 T_o. \tag{12}$$

Thus the Fermi temperature, below which extreme quantum effects occur, does indeed lie in the region where such effects are expected, based upon thermal wavelength considerations—and inspired by the spreading metaphor.

In summary, the view of inhibited spatial spreading as temperature is lowered and thermal wavelength increases is consistent with known behavior of entropy decreasing toward zero with decreasing temperature.

2.6 Free and Non-Free Expansions

The free expansion, which was alluded to earlier, is conceptually simple, but it illustrates profound truths about entropy. In its most basic form, it entails a container with

---

[6]One subtlety is that λ^3 does not represent a rigid volume; it is only an indicator of the range over which quantum interference becomes strong.

[7]S. Chu argued that B-E condensation can be viewed in terms of increased particles wavelength; American Institute of Physics Symposium: Diverse Frontiers of Science, May 3, 2006, Washington, DC. This provided the impetus for addressing low temperature ideal gases in this section.

insulating walls and total volume $2V$. A partition separates the container into left and right chambers, each with volume V. There is a gas in the left chamber and the right chamber is devoid of matter. The insulation is breached briefly while the gas temperature T is measured. The thermometer is removed and the insulation is replaced. The partition is then removed quickly enough that few collisions with gas molecules occur during the process; i.e., removal does not disturb the gas measurably and the internal energy change of the gas is zero: $\Delta E = 0$. According to the discussion in Sect. 2, the phase space volume has increased and thus the entropy has increased, because gas energy has become spread over a larger spatial volume.

If the gas is ideal and is in the classical region of temperature and density, its temperature does not change. If interparticle forces exist, the temperature can either decrease or increase, depending on its initial temperature [30]. There exists an inversion temperature T_I that is similar to the more well known inversion temperature for Joule throttling processes. For temperatures above T_I, the gas temperature *increases* in a free expansion. At temperatures below T_I, the gas temperature *decreases*. Typically, T_I is higher than is encountered under most experimental conditions and free expansions normally lead to cooling. Helium has $T_I \approx 200$ K, so for a free expansion beginning at room temperature, helium's temperature rises, giving a further contribution to gas entropy's increase. The crux of the matter is this: If the gas particles have sufficiently high kinetic energies, they can get close enough to other particles that the average force is repulsive. In this case, expansion to a larger volume diminishes this effect and reduces the average potential energy, which demands a concomitant average kinetic energy increase to assure that $\Delta E = 0$.

In terms of spreading, when the average force between molecules is repulsive, spatial spreading is *enhanced* and this adds to the effect of the volume increase. Thus larger entropy changes are expected, consistent with an increased volume *plus* a higher temperature. In the more common situation, when the average force between molecules is attractive, spreading is constrained and this reduces the net amount of spatial spreading. This is consistent with the effect of increased volume being partially offset by a temperature decrease. Because temperature changes experienced in free expansions are relatively small, all free expansions lead to a *net* increase in spatial spreading and thus entropy.

To close this section, consider an isothermal expansion of a classical fluid, for which [31, 32],

$$(\Delta S_{\text{nonideal}})_T \geq (\Delta S_{\text{ideal}})_T. \tag{13}$$

A spreading interpretation is that interparticle forces in the nonideal system constrain spreading and depress entropy. In an expansion, the average interparticle distance between gas particles lessens, spreading is less constrained, and ultimately—for large enough volume—approaches the degree of spreading of an ideal gas. The left side of (13) exceeds the right side because of the initial depressed degree of spreading at smaller volume, where interparticle forces are strongest on average.

2.7 Mixing of Gases

Consider two gases, each with N molecules, initially in separate but equal volumes V separated by a central partition, as shown in Fig. 4(a). The entire system has tem-

Fig. 4 **(a)**–**(b)** Process I:
Mixing of two ideal gases (black
and empty circles) under
classical conditions, each
initially in volume V and finally
sharing volume $2V$. **(b)**–**(c)**
Process II: Isothermal
compression to volume V
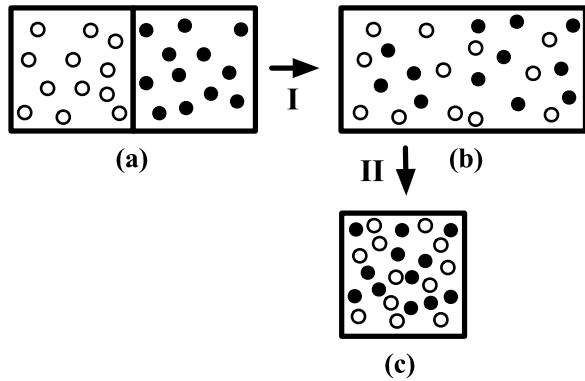
perature T. If the partition is removed, the gases spontaneously mix together, there is
no temperature change, and the entropy change is

$$\Delta S = \begin{cases} 2Nk_B \ln 2 & \text{for distinguishable gases,} \\ 0 & \text{for identical gases.} \end{cases} \quad (14)$$

For *distinguishable* molecules, process I yields the standard "entropy of mix-
ing", (15). Despite its name, this entropy change actually comes from the expan-
sion of each species; i.e., $\Delta S_I = Nk_B \ln 2 + Nk_B \ln 2 = 2Nk_B \ln 2$. Process II,
Fig. 4(b), (c), is an isothermal compression of the container in (b), so $\Delta S_{II} =
-2Nk_B \ln 2$. In the latter process, energy spreading is *negative* for the gas, with pos-
itive energy spreading occurring in the constant-temperature environment, which re-
ceives energy. Thus

$$\Delta S_{\text{total}} = \Delta S_I + \Delta S_{II} = 0 \quad \text{for } \textit{distinguishable} \text{ particles.} \quad (15)$$

Here is an energy spreading interpretation. For *distinguishable* gases, energy
spreads from volume V to $2V$ by *each* species in process I, accounting for $\Delta S_I =
2Nk_B \ln 2$. Note that upon removal of the partition, the energy spectrum of each
species becomes compressed because of the volume increase and, most important,
spreads through the entire container. That is, energy states of "black" particles exist in
the right side as well as the left, with a corresponding statement for "white" particles.
In process II, each species gets compressed from $2V$ to V, so $\Delta S_{II} = -2Nk_B \ln 2$.
In configuration (c), each species spreads energy over volume V, just as in (a), con-
sistent with the overall result $\Delta S_{\text{total}} = 0$.

For *indistinguishable* molecules (black circles become white) in process I there is
no change in energy spreading because the lone species energy was already spreading
energy in both chambers, and $\Delta S_I = 0$.[8] In process II, the lone species is compressed
from $2V$ to V, so

$$\Delta S_{\text{total}} = \Delta S_{II} = -2Nk_B \ln 2 \quad \text{for } \textit{indistinguishable} \text{ particles.} \quad (16)$$

---

[8]In Fig. 4(a), the N particles on the left and N particles on the right have identical energy spectra. In
Fig. 4(b), the $2N$ particles have a single, compressed spectrum because of the volume increase. In this
view, the result in (15) is not apparent.

In (c), $2N$ molecules of the lone species spread energy over volume V, while in (a), N molecules of the species spread energy over volume V and another N molecules spread energy over a *different* volume V. Thus, there is more spreading of the given species and higher entropy in (a), consistent with $\Delta S_{\text{total}} = \Delta S_I + \Delta S_{II} = -2Nk_B \ln 2$.

## 2.8 Metastability, Frozen-in States, and Maxwell's Demon

It has been assumed throughout that internal energy spreading proceeds unimpeded in macroscopic systems, consistent with existing *external* constraints. However, in some cases there exists *internal* metastability, which inhibits full spreading of energy over space and energy modes. This occurs in the cooling of some materials toward 0 K. For example, glycerin can be cooled to low temperatures without it becoming a solid. A consequence is that near absolute zero *supercooled* glycerin's entropy exceeds the value of solid glycerin at the same temperature. For the supercooled "glassy liquid," full energy spreading to the cold surroundings has not occurred and the substance is in an internal metastable state, which could be destroyed—with a concomitant energy transfer to the surroundings—by a small perturbation. This situation is commonly described in terms of the glassy liquid being "more disordered" than the corresponding solid. The latter description hides the important point that energy interactions within the glycerin are too weak to achieve full spreading—i.e., thermal equilibrium—on the time scale of the experiments. Put differently, the lowest possible energy state has not been reached.

An interesting situation exists where states are "frozen in," and in this case, it is used to good advantage. A digital memory maintains a state (e.g., magnetization "upward") because of the existence of strong internal forces. In the simplest possible case, one can envisage a single particle in a double potential well, say with occupation of the left well denoting a "one" and occupation of the right well denoting a "zero." Once a particle is in one of the wells, a central potential barrier prevents passage to the other well under normal temperatures. To erase such a memory one must either lower the central barrier or raise the particle temperature so that the particle has no left-right preference.

Rolf Landauer addressed the issue of memory erasure in 1961 [33] and concluded that erasure of *one bit* at temperature T requires dissipation of at least energy $k_B T \ln 2$ to the environment, with entropy change $\Delta(S_{\text{system}} + S_{\text{surroundings}}) \geq k_B \ln 2$. This result, commonly called *Landauer's Theorem*, has been used to argue that memory erasure "saves" the second law of thermodynamics from the mischief of a Maxwell's demon. The simplest way to understand this is through an ingenious model proposed by Szilard [34–37]. He envisaged the hypothetical one-particle heat engine illustrated in Fig. 5.

Insertion of a partition leaves the particle either on the left or right. With information on which side it resides, one would know how to configure a pulley system and external weight pan to enable the "gas" to do work lifting the weight pan. Once this is done the partition is unlocked, becoming a frictionless, movable piston, and the gas can do external work $W = k_B \ln 2$ as the gas volume is doubled, lifting the weight pan. Strictly speaking, to keep the process slow and reversible, grains of sand

Fig. 5 Szilard one-particle gas engine, as described in the text
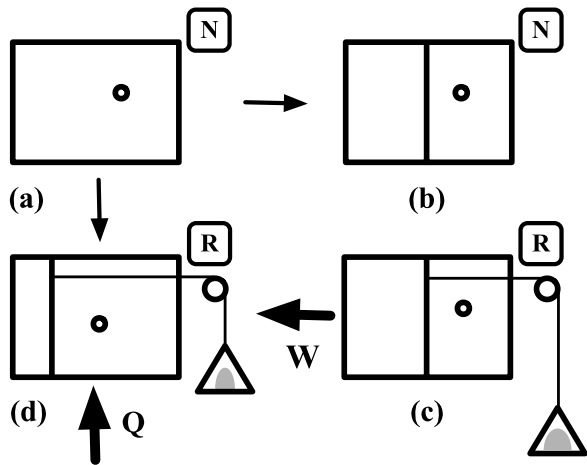
can be removed from the weight pan and deposited on shelves (not shown) as the gas pressure decreases. Replenishment heat energy $Q_{in} = k_B T \ln 2$ would come in from the surroundings, through the container walls and to the particle itself. Measurement of the particle's location after partition insertion could be done in principle by a tiny *intelligent* being, namely, one with great microscopic observation skills, speed, and a memory. This constitutes what has come to be called a Maxwell demon [37].

After this work is done, the pulley system and partition are removed and the gas is in its initial state. A weight has been lifted and the environment has lost some energy because of the energy transfer to the container, and the particle. In essence, heat has been converted to work and, ignoring the demon's memory, it appears that the second law of thermodynamics has been violated, albeit on a microscopic scale. But, as pointed out first by Penrose [38] and independently by Bennett [39], the state of the demon's memory cannot be ignored.

In Fig. 5(a)–(d), the state of the demon's memory (RAM) is shown in the small upper right box. It is initially neutral (N, which can be either R or L), but after measurement, it is R (for right side in this example). After the cycle is run, the memory must be brought back to its initial, neutral state; i.e., its memory must be erased and the initial memory state restored, in order to do complete entropy bookkeeping. According to Landauer's theorem, erasure will send at least energy $Q_{erasure} \geq k_B T \ln 2$ to the surroundings. The latter energy (at least) replenishes the surroundings, and the work needed to effect memory erasure "pays" for the lifted weight's gravitational potential energy increase. Assuming thermodynamics ideas can be sensibly applied at this microscopic level, and that Landauer's theorem holds, the second law is saved.

In terms of the spreading metaphor, without accounting for the demon, the surroundings would have suffered negative energy spreading because of its loss of energy. Because there is no energy spreading associated with ideal frictionless macroscopic work, this signals a net entropy drop for the isolated system of Szilard engine plus surroundings. But after erasure of the memory the situation is very different:

There is a non-negative net change in energy spreading and no violation of the second law.

The most interesting aspect of this problem is the situation *just before* memory erasure. The environment still has less entropy than it had initially, the weight has been lifted, and the memory state is either R or L. The question is: at this point in time, is the second law intact? For this to be so, an entropy of (at least) $k_B \ln 2$ must be attributable to the memory. Because the memory's bit state is "frozen in," there is zero spreading over the R and L states; the memory is either in one or the other.[9] From a spreading viewpoint, the second law is violated just prior to erasure, but is saved by the erasure process.

In contrast, a missing information view leads to assignment of entropy $k_B \ln 2$ based upon the absence of information about the bit state of the memory. Similarly, an algorithmic information view leads to the same entropy for the memory, based upon the number of bits of information needed to specify its state [40]. The dichotomy between the spreading and information theory approaches is notable. Related difficulties have generated criticisms of arguments that the second law can be saved by memory erasure for the Maxwell's demon conundrum [41, 42].

Finally, it is observed that the validity of Landauer's theorem, and indeed the second law itself, has come in question for small or mesoscopic systems under extreme quantum conditions [1, 43–47]. These are beyond the scope of this article.

3 Seeking a Spreading Function

We have argued that entropy is a function that represents a measure of spatial spreading of energy and a temporal spreading over energy states. In this section, we indicate how one can arrive at entropy by seeking a function that is a measure of spreading. The value of this exercise is that it demonstrates the richness and depth of the spreading metaphor. Typically, existing constraints restrict the spreading of energy spatially. Examples of such constraints are an insulating wall, immovable partition, non-porous wall, open electrical switch, and electric or magnetic shielding.

Processes that increase energy spreading are driven by gradients—e.g., of temperature, pressure, or chemical potential. Removal of such a constraint will lead to increased energy spreading that proceeds so as to decrease the gradient(s). It has been said [48] that "nature abhors a gradient," and indeed, all *changes* in the degree of spreading are driven by gradients. Examples are replacing an insulating wall with a conducting one (in the presence of a temperature gradient), making an immovable wall movable (in the presence of a pressure gradient), and removal of a non-porous partition (when there is a chemical potential gradient).

Consider an energy exchange between two *identical* macroscopic bodies that are initially at different temperatures. The higher temperature body has larger internal energy, and the energy exchange allows the energy to become distributed *equitably*. Symmetry dictates that the final internal energies be equal. Evidently, energy spreads

---

[9]It is assumed that the thermodynamic entropy, and concomitant spreading, associated with the memory's other degrees of freedom, have not changed throughout the engine's operation.

over space maximally. If it spreads *less* than maximally, the lower temperature body would end with a temperature below the other body, and if it spread 'too much,' the initially lower temperature body would become the higher temperature one—and such configurations are equivalent to the less than maximal ones that were discussed already.

Does a spreading function exist for this two body system? That is, can energy spreading be quantified? We require that a *bona fide* spreading function, which we call $\mathcal{J}$, have the following properties. In essence, these are postulates that are deemed reasonable for such a function.

1. *For a homogeneous body, $\mathcal{J}$ is a function of the system's energy E, volume V, and particle number N.* Rationale: These are common thermodynamic variables, for one-component systems.
2. *At uniform temperature, $\mathcal{J}$ is an increasing function of energy E.* Rationale: More energy means there is more energy to spread and thus more spreading.
3. *For a body made of the same material, but with twice the volume and energy, the value $\mathcal{J}$ is double that of the original body; i.e., $\mathcal{J}(2E, 2V, 2N) = 2\mathcal{J}(E, V, N)$. This requires the degree of spreading to double when twice the energy spreads over twice the number of particles occupying twice the volume. A generalization is that for any real $\beta \geq 0$, $\mathcal{J}(\beta E, \beta V, \beta N) = \beta \mathcal{J}(E, V, N)$, which is the property called extensivity.*[10] Rationale: Clearly there should be *more* spreading for $\beta > 1$, and it is reasonable that, for example, doubling E, V, N also doubles $\mathcal{J}(E, V, N)$. Halving E, V, N would give half as much spreading, and so forth.
4. *For two systems, labeled a and b, $\mathcal{J}_{a+b} = \mathcal{J}_a(E_a, V_a, N_a) + \mathcal{J}_b(E_b, V_b, N_b)$.* Rationale: We require that the spreading function be additive, as are internal energy, volume, and particle number themselves. This is based on the belief that spreading effects on boundaries of systems *a* and *b* are negligible, being overwhelmed by the spreading over volumes. It is assumed that particles interact via short-range interatomic and intermolecular forces.
5. *$\mathcal{J}_{a+b}$ is maximal at equilibrium.*[11] Rationale: Spreading continues until equilibrium is reached and cannot become larger. This will become more clear in the following example.

Using a symmetry argument, the assumed existence of maximum spatial spreading at equilibrium for *identical* systems demands that if the initial body energies are E_a and $E_b > E_a$, then the final energy of each is

$$E_f = \frac{1}{2}(E_a + E_b). \tag{17}$$

According to postulate 4, after equilibration,

$$\mathcal{J}_{a+b} = \mathcal{J}_a(E_f) + \mathcal{J}_b(E_f) = 2\mathcal{J}(E_f), \tag{18}$$

---

[10]Systems with long-range forces—e.g., gravitational and electric—do *not* satisfy the extensivity condition and the framework here does not apply to such systems.

[11]The development here is for equilibrium states only. It is not known if a generalization of the spreading picture to nonequilibrium states is possible.

Fig. 6 Spreading function
$\mathcal{J}(E, V, N)$ as a function of E
for two identical system with
initial energies E_a and E_b, and
final energy E_f for each

$\mathcal{J}$ vs. E

E_a E_c E_b
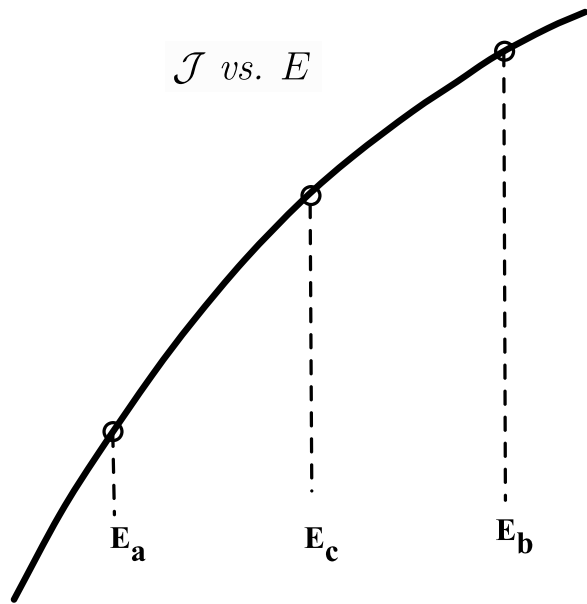
and postulate 5 demands that

$$\mathcal{J}_{a+b}(E_f) = 2\mathcal{J}(E_f) \geq \mathcal{J}(E_a) + \mathcal{J}(E_b) \quad \text{with } E_a \leq E_f \leq E_b. \qquad (19)$$

Expression (19) defines a concave function, and the temperature equilibration process
looks as shown in Fig. 6. In the final state, the value of $\mathcal{J}$ is the same for each body
and, of course, all derivatives are the same. Figure 6 shows that although systems
a and b each experience the same magnitude energy change, concavity assures that
system a has a greater increase in its spreading function than system b.

Now consider two bodies of the same material, one (system a) of which has twice
the volume and particle number of the other (system b), with total energy $E_a + E_b =
E$. Suppose the two bodies each have equal initial energies: $E_{a,i} = E_{b,i} = E/2$. We
expect system a's initial temperature to be lower than system b's, given that is has
the same energy as the smaller system b. When the two systems are put in thermal
contact with one another, energy will spread until it is shared *equitably* by the two
subsystems, namely, with final energies $E_a = 2E/3$ and $E_b = E/3$. A bit of mathe-
matics shows that this is consistent with $\mathcal{J}_{a+b}$ being maximized; i.e., *equitable* energy
sharing occurs when the total spreading function is maximized relative to the existing
constraint $E_a + E_b = E$. Specifically,

$$(\partial\mathcal{J}_a/\partial E_a)|_{E_a=2E/3} = (\partial\mathcal{J}_b/\partial E_b)|_{E_b=E/3}. \qquad (20)$$

This analysis can easily be extended to system a being β times larger than
the otherwise identical system b. After equilibration, the final energies will be
$E_{a,f} = \beta E/(\beta + 1)$ and $E_{b,f} = E/(\beta + 1)$, with $(\partial\mathcal{J}_a/\partial E_a)|_{E_{a,f}=\beta E/(\beta+1)} =
(\partial\mathcal{J}_b/\partial E_b)|_{E_{b,f}=E/(\beta+1)}$. Again, this maximizes J_{a+b}. More generally, for *any* two
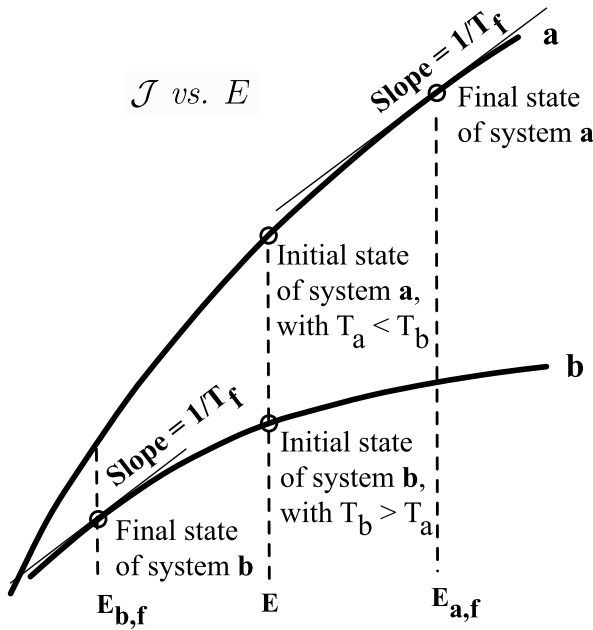
Fig. 7 Spreading functions $\mathcal{J}(E, V, N)$ *vs. E* for two different materials. Initially, the systems have the same energy E, and $(\partial \mathcal{J}_a/\partial E_a)|_{E_a=E} > (\partial \mathcal{J}_b/\partial E_b)|_{E_b=E}$. Concavity guarantees that the *only* energies for which an infinitesimal variation of E_a and E_b with $E_a + E_b$ fixed gives $d\mathcal{J} = 0$ are $E_{a,f}$ and $E_{b,f}$, at which $(\partial \mathcal{J}_a/\partial E_a)|_{E_a=E_{a,f}} = (\partial \mathcal{J}_b/\partial E)|_{E_b=E_{b,f}}$

different systems, of different types and/or sizes, maximization of spreading leads to the following generalization of (20),

$$(\partial \mathcal{J}_a/\partial E_a)|_{E_a=E_{a,f}} = (\partial \mathcal{J}_b/\partial E_b)|_{E_b=E_{b,f}} \tag{21}$$

which determines $E_{a,f}$ and $E_{b,f} = E - E_{a,f}$. Here, we assume, based on the pattern seen above, that *equitable* energy sharing occurs when spreading is maximized.

Figure 7 shows how the spreading functions and the equilibration process might look for two different-sized and/or different type systems that begin with equal energies and then share energy until thermal equilibrium exists. Given the above discussion and the observation, using Fig. 7, that $(\partial \mathcal{J}_a/\partial E_a)|_{E_a=E} > (\partial \mathcal{J}_b/\partial E_b)|_{E_b=E}$ it is suggestive that $\partial \mathcal{J}/\partial E$ is inversely related to temperature. The simplest such relationship is

$$(\partial \mathcal{J}/\partial E)_{V,N} = 1/T. \tag{22}$$

With this definition, $\partial \mathcal{J}/\partial E$ decreases and T increases as energy increases for each system. Furthermore, $\mathcal{J}_a$ increases more than $\mathcal{J}_b$ decreases, consistent with the total spreading function $\mathcal{J} = \mathcal{J}_a + \mathcal{J}_b$ increasing to a maximum at equilibrium.

For a constant-volume heating process that proceeds along a given $\mathcal{J}$ curve, $dE = \delta Q$, where δQ is the (inexact) heat differential. Equation (22) implies that $d\mathcal{J} = dE/T \equiv \delta Q/T$, in analogy with the Clausius entropy form $dS = \delta Q/T$.

Thus, with the temperature definition (22), the spreading function $\mathcal{J}$ shares the important mathematical property $d\mathcal{J} = \delta Q/T$ with entropy S. More connections between $\mathcal{J}(E, V, N)$ and $S(E, V, N)$ can be made, and the interested reader is directed to Ref. [19] for further details. The main conclusion is

$$\mathcal{J} \longleftrightarrow S. \tag{23}$$

Entropy has properties of a spreading function and, in turn, one can use expected properties of a spreading function to obtain the entropy function.

4 Concluding Remarks and Questions

Although entropy is intimately related to the second law of thermodynamics, it is also true that the well known statements of the second law by Clausius and by Kelvin and Planck neither use nor require the entropy concept. Čápek and Sheehan [1] have pointed out that most current challenges to the second law entail heat and work, rather than entropy. Further, they wrote, "Entropy remains enigmatic. The more closely one studies it, the less clear it becomes. Like a pointillism painting whose meaning dissolves into a collection of meaningless points when observed too closely, so too entropy begins to lose meaning when one contemplates it at a microscopic level."

The scope of this article is highly limited, being directed primarily at developing the spreading metaphor as an *interpretive* tool. It is by no means clear whether the spreading concept can developed more fully mathematically, for example, for systems not in thermodynamic equilibrium. Were that possible, one might hope for a way to address non-equilibrium entropy. Similarly it is not clear if the spreading concept can be usefully extended to non-extensive systems characterized by long-range gravitational and/or electric forces.

Questions abound. Can improved language help to clarify the meaning and improve the utility of entropy and, if so, can the spreading metaphor help in this regard? Can the spreading metaphor be helpful when S is not well defined? Can this metaphor shed light on situations where the second law of thermodynamics is suspected of being, or is shown to be, violated? Can the spreading concept be used in combination with one or more other metaphors—e.g., multiplicity and/or missing information—to provide a more complete qualitative description of entropy? It is hoped that time will bring answers to these questions.

References

1. Čápek, V., Sheehan, D.P.: Challenges to the Second Law of Thermodynamics—Theory and Experiment. Springer, Dordrecht (2005)
2. Boltzmann, L.: Lectures on Gas Theory—Part I, pp. 441–443. Barth, Leipzig (1896), reprinted by University of California Press, Berkeley (1964)
3. Burgers, J.M.: Entropy and disorder. Br. J. Philos. Sci. **5**, 70–71 (1954)

4. Styer, D.F.: Insight into entropy. Am. J. Phys. **68**, 1090–1096 (2000)
5. Darrow, K.K.: The concept of entropy. Am. J. Phys. **12**, 183–196 (1944)
6. Bohren, C.F., Albrecht, B.A.: Atmospheric Thermodynamics, pp. 150–152. Oxford University Press, Oxford (1998)
7. Dingle, H.: Bull. Inst. Phys. **10**, 218 (1959)
8. Wright, P.G.: Entropy and disorder. Contemp. Phys. **11**, 581–588 (1970)
9. Lambert, F.L.: Disorder—a cracked crutch for supporting entropy discussions. J. Chem. Ed. **79**, 187–192 (2002), revised web version: http://www.entropysite.com/cracked_crutch.html. See also Lambert's website, http://www.entropysite.com for related discussions
10. Gibbs, J.W.: The Scientific Papers of J. Willard Gibbs, vol. 1, p. 418. Constable, London (1906), reprinted by Dover, New York (1961)
11. Jaynes, E.T.: Information theory and statistical mechanics. Phys. Rev. **106**, 620–630 (1957)
12. Baierlein, R.: Teaching the approach to thermodynamic equilibrium: some pictures that help. Am. J. Phys. **46**, 1042–1045 (1978)
13. Swanson, R.M.: Entropy measures amount of choice. J. Chem. Ed. **67**, 206–208 (1990)
14. Brissaud, J.B.: The meanings of entropy. Entropy **7**, 68–96 (2005)
15. Clausius, R.: On the application of the theorem of the equivalence of transformations to the internal work of a mass of matter. Philos. Mag. **24**, 81–97 (1862), Reprinted in Kestin, J. (ed.) The Second Law of Thermodynamics. Dowden, Hutchinson & Ross, Stroudsberg (1976)
16. Klein, M.J.: Gibbs on Clausius. Hist. Stud. Phys. Sci. **1**, 127–149 (1969)
17. Denbigh, K.G.: The many faces of irreversibility. Br. J. Philos. Sci. **40**, 501–518 (1989)
18. Denbigh, K.G.: The Principles of Chemical Equilibrium. Cambridge University Press, Cambridge (1961), Sect. 1.7
19. Leff, H.S.: Thermodynamic entropy: the spreading and sharing of energy. Am. J. Phys. **64**, 1261–1271 (1996)
20. Lambert, F.L.: Entropy is simple, qualitatively. J. Chem. Ed. **79**, 1241–1246 (2002), revised web version: http://www.entropysite.com/entropy_is_simple/index.html
21. Gross, D.H.E.: Microcanonical Thermodynamics: Phase Transitions in 'Small' Systems. World Scientific, Singapore (2001)
22. Leff, H.S.: What if entropy were dimensionless? Am. J. Phys. **67**, 1114–1122 (1999)
23. Falk, H.: Physica **29**, 1114 (1963)
24. Baierlein, R.: Forces, uncertainty, and the Gibbs entropy. Am. J. Phys. **36**, 625–629 (1968)
25. Leff, H.S.: Entropy differences between ideal and nonideal systems. Am. J. Phys. **37**, 548–553 (1969)
26. Bruch, L.W., Schick, M., Siddon, R.L.: Limitation on the quantum-mechanical extension of Baierlein's entropy theorem. Am. J. Phys. **44**, 1007–1008 (1976)
27. Leff, H.S.: Teaching the photon gas in introductory physics. Am. J. Phys. **70**, 792–797 (2002)
28. Anderson, M.H., Ensher, J.R., Matthews, M.R., Wieman, C.E., Cornell, E.A.: Observation of Bose–Einstein condensation in a dilute atomic vapor. Science **269**, 198–201 (1995)
29. Pathria, R.K.: Statistical Mechanics, 2nd edn. Butterworth–Heinemann, Oxford (1996), Chaps. 7, 8
30. Goussard, J., Roulet, B.: Free expansion for real gases. Am. J. Phys. **61**, 845–848 (1993)
31. Lesk, A.M.: Entropy changes in isothermal expansions of real gases. Am. J. Phys. **42**, 1030–1033 (1974)
32. Leff, H.S.: Entropy changes in real gases and liquids. Am. J. Phys. **43**, 1098–1100 (1975)
33. Landauer, R.: Irreversibility and heat generation in the computing process. IBM J. Res. Dev. **5**, 183–191 (1961)
34. Szilard, L.: On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. Z. Phys. **53**, 840–856 (1929)
35. Feld, B.T., Weiss Szilard, G.: The Collected Works of Leo Szilard: Scientific Papers. pp. 103–129. MIT Press, Cambridge (1972), English translation is by A. Rapoport and M. Knoller
36. Wheeler, J.A., Zurek, W.H.: Quantum Theory and Measurement, pp. 539–548. Princeton University Press, Princeton (1983)
37. Leff, H.S., Rex, A.F. (eds.): Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing. Institute of Physics, Bristol (2003), now distributed by Taylor & Francis
38. Penrose, O.: Foundations of Statistical Mechanics, pp. 221–238. Pergamon, Oxford (1970)
39. Bennett, C.H.: The thermodynamics of computation—a review. Int. J. Theor. Phys. **21**, 905–940 (1982), reprinted in Ref. [37]
40. Zurek, W.H.: Algorithmic randomness and physical entropy. Phys. Rev. **40**, 4731–4751 (1989)
41. Earman, J., Norton, J.D.: Exorcist XIV: the wrath of Maxwell's demon, part I: from Maxwell to Szilard. Studies Hist. Philos. Mod. Phys. **29**, 435–471 (1998)

42. Earman, J., Norton, J.D.: Exorcist XIV: the wrath of Maxwell's demon, part II: from Szilard to Landauer and beyond. Studies Hist. Philos. Mod. Phys. **30**, 1–40 (1999)
43. Allahverdyan, A.E., Nieuwenhuizen, T.M.: Breakdown of the Landauer bound for information erasure in the quantum regime. Phys. Rev. E **64**, 0561171–0561179 (2001)
44. Allahverdyan, A.E., Nieuwenhuizen, T.M.: Testing the violation of the Clausius inequality in nanoscale electric circuits. Phys. Rev. B **66**, 1153091–1153095 (2002)
45. Allahverdyan, A.E., Balian, R., Nieuwenhuizen, T.M.: Thomson's formulation of the second law: an exact theorem and limits of its validity. In: Sheehan, D.P. (ed.) Quantum Limits to the Second Law. AIP Conference Proceedings, vol. 643. American Institute of Physics, New York (2002)
46. Allahverdyan, A.E., Nieuwenhuizen, T.M.: Unmasking Maxwell's demon. In: Sheehan, D.P. (ed.) Quantum Limits to the Second Law. AIP Conference Proceedings, vol. 643. American Institute of Physics, New York (2002)
47. Nieuwenhuizen, T.M., Allahverdyan, A.E.: Quantum Brownian motion and its conflict with the second law. In: Sheehan, D.P. (ed.) Quantum Limits to the Second Law. AIP Conference Proceedings, vol. 643. American Institute of Physics, New York (2002)
48. Schneider, E.D., Sagan, D.: Into the Cool: Energy Flow, Thermodynamics, and Life. University of Chicago Press, Chicago (2005)

How physicists disagree on the meaning of entropy

Robert H. Swendsen[a)]
Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Discussions of the foundations of statistical mechanics, how they lead to thermodynamics, and the appropriate definition of entropy have occasioned many disagreements. I believe that some or all of these disagreements arise from differing, but unstated assumptions, which can make opposing opinions difficult to reconcile. To make these assumptions explicit, I discuss the principles that have guided my own thinking about the foundations of statistical mechanics, the microscopic origins of thermodynamics, and the definition of entropy. The purpose of this paper will be fulfilled if it paves the way to a final consensus, whether or not that consensus agrees with my point of view. © *2011 American Association of Physics Teachers.*
[DOI: 10.1119/1.3536633]

I. INTRODUCTION

> "Nobody really knows what entropy really is."
> —John von Neumann[1]

Since I began speaking and publishing on the relation between statistical mechanics and thermodynamics in general and the meaning of entropy in particular,[2–7] I've encountered a diversity of opinion among experts that is remarkable for a field that is well over a century old. Most colleagues with whom I have discussed the matter have indicated that they believe their views are essentially the same as those of the majority of physicists. However, when we discuss details, opinions turn out to be quite diverse and, at times, contentious.[4,8]

The following is a partial list of opinions I have encountered in the literature and in discussions with other scientists:

- The theory of probability has nothing to do with statistical mechanics.
- The theory of probability is the basis of statistical mechanics.
- The entropy of an ideal classical gas of distinguishable particles is not extensive.
- The entropy of an ideal classical gas of distinguishable particles is extensive.
- The properties of macroscopic classical systems with distinguishable and indistinguishable particles are different.
- The properties of macroscopic classical systems with distinguishable and indistinguishable particles are the same.
- The entropy of a classical ideal gas of distinguishable particles is not additive.
- The entropy of a classical ideal gas of distinguishable particles is additive.
- Boltzmann defined the entropy of a classical system by the logarithm of a volume in phase space.
- Boltzmann did not define the entropy by the logarithm of a volume in phase space.
- The symbol W in the equation $S = k \log W$, which is inscribed on Boltzmann's tombstone, refers to a volume in phase space.
- The symbol W in the equation $S = k \log W$, which is inscribed on Boltzmann's tombstone, refers to the German word "Wahrscheinlichkeit" (probability).
- The entropy should be defined in terms of the properties of an isolated system.
- The entropy should be defined in terms of the properties of a composite system.
- Thermodynamics is only valid in the "thermodynamic limit," that is, in the limit of infinite system size.
- Thermodynamics is valid for finite systems.
- Extensivity is essential to thermodynamics.
- Extensivity is not essential to thermodynamics.

This remarkable diversity of opinion has an interesting consequence. When people discuss the foundations of statistical mechanics, the justification of thermodynamics, or the meaning of entropy, they tend to assume that the basic principles they hold are shared by others. These principles often go unspoken, because they are regarded as obvious. It has occurred to me that it might be good to restart the discussion of these issues by stating basic assumptions clearly and explicitly, no matter how obvious they might seem. This paper is a start in that direction.

There are two possible reactions to the principles I put forward. A reader might agree with them. In that case, we would have a firm basis on which to proceed. Or, a reader might take issue with one or more. In that case, we would know where the conflict lies, which would give us a good chance of resolving points of disagreement. In either case, we should be able to make progress toward arriving at a consensus, which is the goal of this paper.

Because my topic is limited to macroscopic measurements of macroscopic systems, I will discuss what I understand those terms to mean in Sec. II. In this paper I will put forward 12 principles based on the concept of macroscopic measurements that have led me to advocate the use of Boltzmann's 1877 definition of the entropy[11] over other definitions that are often found in textbooks.

II. MACROSCOPIC SYSTEMS

In this paper I am concerned with the question of how to describe the observed behavior of macroscopic systems. The concept of macroscopic frames all of my arguments, so it is important to make clear at the outset how I define it. A macroscopic system contains a large number of particles, and a macroscopic measurement is limited in its resolution. These two features are closely related, in that what can be regarded as a large number depends on the resolution of the macroscopic measurements.

The reason for specifying a large number of particles is

that the quantities of interest in thermodynamics are collective variables, such as the energy or the number of particles in a system. The *relative* statistical fluctuations of such quantities are generally inversely proportional to the square root of the number of particles. If the statistical fluctuations are much smaller than the resolution of the macroscopic measurements, they can be ignored; the average values obtained from statistical mechanics then give a description of the expected results of the experiment.

In the late 19th century, when Boltzmann and Gibbs did their seminal work, the existence of atoms had not been proven. The idea of experimentally observing atomic behavior was scarcely considered. Now, it is possible to obtain images of microscopic structure with atomic resolution. Nevertheless, I am restricting my attention in this paper to macroscopic measurements that cannot discern microscopic behavior in order to discuss the emergence of a thermodynamic description from statistical mechanics.

In the remainder of the paper I will give the rationale for each of the principles I have followed, and I will show how they lead to the adoption of Boltzmann's 1877 definition of the entropy in terms of the logarithm of the probability of macroscopic states for a composite system.

III. PROBABILITY OF MACROSCOPIC STATES

Principle 1: Probability theory is necessary for a theoretical description of macroscopic behavior.

The first—and most fundamental—principle is that the basis for obtaining a description of a macroscopic system from microscopic laws of motion is given by probability theory. In any experiment (real or gedanken), the system is in some specific microscopic state (quantum or classical) at any given instant. That microscopic state is a property of the system, independent of measurement.

The most immediate consequence of the limited resolution of macroscopic measurements is that it severely restricts our knowledge of the microscopic state of a system. We cannot determine the microscopic state experimentally—we can only eliminate microscopic states that are not consistent with our macroscopic observations.

The limitations on our knowledge bring us to the distinction between reality and our knowledge of reality. The reality is the microscopic state of the system at any given time. Our knowledge of reality consists of the information we obtain from macroscopic measurements and the conclusions we are able to draw from that information. We can only construct a representation or description of the behavior of the system; we cannot know the microscopic state of a system from macroscopic measurements.

In quantum systems our knowledge is even more limited. For example, except for eigenstates, which have probability zero, the energy is not even determined uniquely by the microscopic state, so it cannot be a property of the system independent of measurement.

The most useful method I know for describing limited knowledge is Bayesian probability theory,[9] which led me to the first principle.

After deciding to use probability theory, there remains the choice of which probability distribution to use. The most reasonable choice would seem to be the simplest that is consistent with what we know from macroscopic observations. Therefore, I take the probability distribution (a Baysian prior or model probability) to be uniform in phase space for iso- lated classical systems (subject to constraints on the total energy and the restriction of the particles to certain volumes), and correspondingly uniform over microscopic states of quantum systems. The logical consequences of such probability distributions are known to lead to predictions that agree with experiment, which is comforting.

Principle 2: Probability theory is sufficient for a theoretical description of macroscopic states.

In one sense, the introduction of probability distributions very nearly completes the theory of many-body systems. Little else is essential. 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.

This principle is very important because it implies that however we define concepts like entropy and free energy in statistical mechanics, the consequences of the definitions must be consistent with the predictions of probability theory if they are to have the properties required by thermodynamics.

IV. COMPOSITE SYSTEMS

Principle 3: Statistical mechanics and thermodynamics must predict the properties of composite systems.

An essential part of statistical mechanics and thermodynamics is the analysis and prediction of the behavior of composite systems. A simple isolated system in equilibrium does not do anything macroscopically measurable. You can't even make an experimental determination of its temperature without putting a thermometer in contact with it, and then you have a composite system.

A simple container full of gas must also be regarded as a composite system if we want to investigate questions such as whether the density of the gas is uniform. Without conceptually dividing the system into smaller subsystems, we cannot discuss density variations.

An important feature of a composite system is that it can have internal constraints between its subsystems. The release of internal constraints can lead to measurable changes, which can be predicted by statistical mechanics and thermodynamics.

Although I don't expect serious disagreement on this principle, it does lead to a different emphasis than the usual textbook discussion. It is common to define thermodynamic functions for isolated systems and only much later consider equilibrium in composite systems. I believe that because of the crucial importance of composite systems, they should play a leading role in the development of statistical mechanics and thermodynamics.

Section V will discuss the measurement of extensive parameters, which are quantities that are proportional to how much of something there is in a system. Examples include the energy and the number of particles. The prediction of the measured values of extensive parameters is a key step in linking statistical mechanics to thermodynamics.

V. PREDICTIONS OF THERMODYNAMIC QUANTITIES

Principle 4: The values of extensive parameters that maximize the probability predict the results of measurements of those parameters for composite systems in equilibrium.

This principle provides the key link between statistical mechanics and thermodynamic measurements.

When a constraint in a composite system is released, measurable quantities can change. As an example, consider a composite system consisting of two subvolumes separated by a partition, each containing some amount of the same type of ideal gas. Each subvolume contains on the order of 10^{20} particles, and our measurement apparatus can resolve the density of the gas to an accuracy of about 10^{-5}. If a hole is punched in the partition, the density of the gas in each subvolume will go to approximately the same value, within fluctuations of the order of 10^{-10}. Because the fluctuations are much smaller than the resolution of our measurement apparatus, we can take the location of the maximum of the probability distribution to predict the experimental outcome. This feature strongly supports Principle 2; probability theory is sufficient to predict macroscopic behavior.

Similar examples can be given for releasing constraints on the energy (using walls that conduct heat) or volume (using a freely moving piston to separate the subvolumes). In each case the probability distribution is very narrow, so that the fluctuations cannot be observed by macroscopic measurements. The extremely small relative fluctuations of macroscopic observables are so universal that, in the 19th century, many of Boltzmann's opponents didn't believe in their existence.

Although nonequilibrium behavior after the release of constraints is both interesting and important, the discussion here is limited to equilibrium states, which are discussed in Sec. VI.

VI. EQUILIBRIUM

Principle 5: A macroscopic equilibrium state is defined by two properties: the probability of macroscopically observable changes is extremely small, and there is no macroscopically observable flux of energy or particles. (This property distinguishes equilibrium from steady state.)

There might be some disagreement on this point. There is a substantial literature in statistical mechanics that makes the fundamental assertion that equilibrium is defined by a particular "equilibrium probability distribution" in phase space (or Hilbert space).

In my opinion, such a view is a serious error, primarily because the probability distribution of the microscopic states is not macroscopically observable. We use probability theory because we cannot discern microscopic states; we certainly cannot measure the relative frequency with which they occur.

If we limit the definition of equilibrium to behavior that can be observed, it follows that there are many probability distributions that all make the same predictions.[10] The simplest probability distribution is the uniform distribution, but it is not unique.

It is traditional to define a number of thermodynamic functions to facilitate the analysis of macroscopic systems in equilibrium. Although Principle 2 implies that these functions are not absolutely necessary, they are such convenient descriptions of macroscopic behavior that it would be unreasonable to do without them. Their general nature is discussed in Sec. VII.

VII. THERMODYNAMIC PREDICTIONS

Principle 6: The predictions of statistical mechanics and thermodynamics are representations or descriptions of a system based on the extent of our knowledge.

This principle again reflects the distinction between reality and our knowledge of reality, between properties of a system and a description or representation of measurable quantities based on our limited knowledge.

As an example of this distinction, consider again a composite system consisting of a box containing a gas, with a partition dividing the box into two equal subvolumes. The partition has a small hole in it, so that molecules of the gas can move between the two subvolumes. At any instant of time, there is some specific number of particles on each side of the partition. Thermodynamics predicts a number of particles that give the same density on both sides of the partition. The predicted number turns out to agree with experiment to within the limited resolution of macroscopic measurements. For this reason, thermodynamics provides a very useful description of the behavior of a macroscopic system.

In contrast, the actual number of particles on each side of the partition at any instant cannot be the number that is predicted. The actual number is not determined for quantum systems without measurement, and even for classical systems, it fluctuates with time. The predicted number is a description based on our knowledge and is constant in time. It is very useful for human purposes, but it is not a real property of the system.

It is sometimes claimed that the predicted number of molecules in each subvolume is a real property of the system if we regard it as an average over the course of an experiment. How long would the observation time have to be for such a claim to be true? Consider an open system with about 10^{20} particles in equilibrium and a corresponding statistical uncertainty of about 10^{10} particles. To reduce the statistical uncertainty of the mean to about one particle, we would need at least 10^{20} independent observations. If the correlation time for the system is about 1 ms, this would take 10^{17} s, which is comparable to the age of the universe. Even with such a long observation time, we would still not have an exact result because the average number of particles is generally not an integer. For any reasonable experiment during the lifetime of a physicist, the prediction of thermodynamics is in error by an enormous number of particles and should not be confused with the actual number of particles.

For the same reasons, the energy, the entropy, and the associated free energies are thermodynamic descriptions rather than real properties of a macroscopic system. The entropy is actually defined at a higher level of abstraction than the energy or the number of particles. That is the subject of Sec. VIII.

The distinction between real properties of a system and our knowledge of the system might seem philosophical and a bit pedantic, but it greatly clarifies some issues that might otherwise be rather puzzling.

VIII. ENTROPY

This section considers the controversial question of what "entropy" means and how to define it. Principle 7 is based on the most important of the thermodynamic properties of the

entropy,[2,12] which leads to both the thermodynamic conditions for equilibrium and the second law of thermodynamics.

Principle 7: The primary property of the entropy is that it is maximized in equilibrium.

Because the macroscopically observable behavior of an isolated system in equilibrium does not change with time, the maximization of the entropy cannot be applied to a simple system. It can be applied to a composite system: simply release a constraint and see what happens. If the definition of the entropy is correct, the location of the maximum of the entropy should predict the observed equilibrium values of extensive macroscopic observables.

Principle 7 also leads directly to the second law of thermodynamics. If the entropy is always maximized in equilibrium for a composite system, then the change in entropy after a constraint is released cannot be negative.

If we compare Principle 7 with the predictions of probability theory, we see that the location of the maximum of the entropy must always coincide with the location of the maximum of the probability distribution.

An immediate consequence of Boltzmann's 1877 definition of the entropy as the logarithm of the probability distribution for macroscopic observables is that the location of the maximum of the entropy *always* coincides with the equilibrium values of those macroscopic observables. If any other definition is used, it requires a separate demonstration to show that it also predicts these values correctly.

The automatic agreement of the predictions of Boltzmann's definition of the entropy with the correct equilibrium values of macroscopic observables makes it the natural choice. It might be possible to define the entropy differently, but the predictions of any alternative definition must be identical to those of Boltzmann's definition in terms of the logarithm of the probability.

Principle 7 completes what I regard as a convincing argument in favor of Boltzmann's 1877 definition of the entropy.

The remainder of the paper takes up issues that are associated with the concept of entropy. Their purpose is both to introduce the remaining principles that have guided my thinking on these issues and to complete the picture presented so far.

IX. ADDITIVITY

Principle 8: Additivity is essential to any consistent definition of the entropy of a system with short-ranged interactions between its particles.

In thermodynamics it is generally assumed that the entropy of a composite system is given by the sum of the entropies of the subsystems. This property is known as "additivity."

For Boltzmann's 1877 definition of the entropy, the validity of the assumption of additivity is based on the short range of molecular interactions, which is much smaller than the dimensions of the system. Only a very small fraction of the particles in one subsystem interact with those in another subsystem, so that the sum of all such interaction energies is still relatively small. If the direct interactions between subsystems can be neglected, the entropy satisfies additivity.

As an aside, using Boltzmann's definition of the entropy suggests the alternative of referring to this property as "separability," because the entropy of a composite system is defined first.

If we were to use a definition of the entropy that did not satisfy additivity and nevertheless wanted to have correct results for composite systems, we could assign an arbitrary function—or simply the value zero—as the entropy of any subsystem. The entropy of a composite system could then be obtained by adding an extra term to recover the Boltzmann expression. It is possible to create such a formalism, but none of the usual expressions for temperature, pressure, or chemical potential in terms of partial derivatives of the entropy would be necessarily valid. Without additivity, we would not have thermodynamics as we know it.

The importance of additivity probably would go without saying if it were not for a suggestion that an otherwise incorrect definition of the entropy might be saved by an extra term for composite systems.[8] I don't see any virtue to such a procedure, and I stand by Principle 8.

X. THE THERMODYNAMIC LIMIT

The thermodynamic limit is defined as the infinite-size limit of the ratios of extensive quantities—ratios such as the energy per particle U/N or the particle density N/V. The advantage of taking the limit of infinite size is that uncertainties in these ratios go to zero because the relative fluctuations are generally proportional to $1/\sqrt{N}$.

Principle 9: The thermodynamic limit is not required for the validity of thermodynamics.

To judge from some textbooks, this principle might be the most controversial of the ones discussed in this paper.

However, the thermodynamic limit is misnamed. It is not essential to the foundations of thermodynamics. It cannot be essential if we are to apply thermodynamics to real systems, which are necessarily finite. We never do experiments on infinite systems. If thermodynamics worked only for infinite systems, it might still be interesting as mathematics, but it would be irrelevant as science.

The thermodynamic limit is mathematically convenient for certain problems. Phase transitions, for example, only exhibit nonanalytic behavior in the thermodynamic limit, which makes for a much cleaner mathematical description. Nevertheless, the thermodynamic limit should not play any essential role in the foundations of statistical mechanics and thermodynamics.

XI. DISTINGUISHABILITY AND INDISTINGUISHABILITY

Principle 10: "Indistinguishability" is a property of microscopic states. It does not depend on experimental resolution.

In my opinion, this principle should be an obvious consequence of the definitions found in any textbook on quantum mechanics. However, I have had enough arguments about it to know that it is far from obvious.

The definitions of distinguishability and indistinguishability are simple: (1) If the exchange of two particles in a system results in a different microscopic state, the particles are distinguishable. (2) If the exchange of two particles in a system results in the original microscopic state, the particles are indistinguishable. (For fermions, two states are usually regarded as identical if they differ only by an overall minus sign.)

The definition of indistinguishability does not have anything to do with the interactions between particles. It is possible in either quantum or classical physics for two distinct

states to have the same energy. Nevertheless, if the microscopic state is different after the exchange of two particles, those particles are distinguishable.

Unfortunately, "distinguishable" is sometimes confused with what might be called "observably different." Two particles are observably different if exchanging them alters the properties of the system in a way that is observable. Clearly, if particles are observably different, they must be distinguishable. In contrast, particles can be distinguishable without their exchange producing any observable differences.

A simple example of this distinction is provided by a mixture of ^{3}He and ^{4}He. It would not be possible for a macroscopic measurement to detect the difference in the microscopic states that would result from exchanging a single ^{3}He atom with a single ^{4}He atom. Nevertheless, there would be a difference in the microscopic states, and the two isotopes of helium are *not* mutually indistinguishable.

The term "identical particles" is often used as a synonym for indistinguishable particles. This use has the unfortunate consequence that a model of classical distinguishable particles with identical properties might be mistaken for a model of indistinguishable particles.

The concept of indistinguishability is foreign to classical mechanics. Consider the trajectory of an isolated classical system in phase space (the $6N$-dimensional space defined by the positions and momenta of all particles in the system) in which the microscopic state is described by a point. If two particles are exchanged at a given time, the trajectory becomes discontinuous. The exchange of particles has resulted in a different microscopic state, regardless of whether the Hamiltonian gives the same energy for the two microstates.

In quantum mechanics N-particle states of indistinguishable particles are characterized by a wave function that has been symmetrized (or antisymmetrized) by summing over all permutations of the particles, with a change in sign for each permutation for fermions, or without a change in sign for bosons.

A classical system of indistinguishable particles can be described by the same procedure. The microscopic state of a classical system of indistinguishable particles would be described by the $N!$ points in phase space found from the set of all permutations of the particles. The trajectory (or trajectories) of the set of $N!$ points is clearly unaffected by the exchange of any two particles at any point in time.

The idea of representing a classical state by $N!$ points in phase space is a bit odd, but that is because indistinguishability is not a classical concept. However, if indistinguishability is to be imposed on a classical system, this representation seems to be the most reasonable way of doing it.

Many textbooks claim that classical systems with distinguishable and indistinguishable particles are described by different expressions for the entropy. However, it is straightforward to demonstrate that the macroscopic properties of a classical system are exactly the same whether the particles are distinguishable or indistinguishable.[2] Since the macroscopic behaviors of classical systems with distinguishable and indistinguishable particles are the same, it seems natural that their entropies should also be the same, which leads to my next principle.

Principle 11: Systems with identical macroscopic properties should be described by the same entropy.

Boltzmann's 1877 definition of the entropy gives the same expression for the entropy for classical systems with either distinguishable or indistinguishable particles.[2] The traditional definition in terms of a volume in phase space, which is often erroneously attributed to Boltzmann,[5] gives different expressions, at least one of which must clearly be incorrect. The worst failings of the traditional definition of the entropy for a system of distinguishable particles are that it violates the second law of thermodynamics and makes incorrect predictions for equilibrium with respect to the exchange of particles between subsystems.[2]

The error in the traditional definition of the entropy of a classical system of distinguishable particles also has the consequence that it predicts that the entropy of an ideal gas is not extensive. This problem is not really fundamental, but it has bothered people. And it leads to the next principle.

XII. EXTENSIVITY

Principle 12: Extensivity is not essential to thermodynamics.

Extensivity is the property that the macroscopic observables of a system are all directly proportional to its size. This property implies that ratios, such as U/N, V/N, and S/N, are all independent of the size of the system. In many textbooks, extensivity is taken to be a fundamental postulate of thermodynamics.[12] It is certainly convenient mathematically, because it leads directly to the Euler and Gibbs–Duhem equations. It is an appropriate assumption when the physical properties of a material are being investigated, and the surface or interface contributions can be neglected.

However, real systems have surfaces and interfaces, which are important topics of research. Because the surface-to-volume ratio changes with the size of the system, real systems are not extensive, and the deviations from extensivity can be very important. For example, a real gas in a real container will usually be adsorbed to some extent on the inner walls of the container. At low temperatures, the fraction of adsorbed molecules can be quite large, which is exploited in the construction of cryopumps.

To describe the thermodynamics of a surface, we must be able to describe the thermodynamics of a nonextensive system and extract the parts of the free energy, etc. that are not directly proportional to the size. Therefore, statistical mechanics and thermodynamics must be applicable to nonextensive systems.

Recognizing that extensivity is not an essential property of thermodynamic systems is important in deciding on an appropriate definition of entropy. Some colleagues claim that a definition of entropy that gives a demonstrably incorrect expression can be made acceptable by imposing extensivity with an additional term of the form $-k_B \ln(N!)$. However, because thermodynamics should also correctly describe nonextensive systems, that is, systems with entropies that cannot be made extensive by a term that depends only on N, such a correction is not feasible.

There is also another difficulty in trying to impose extensivity on the fundamental definition of the entropy. If the system under consideration contains more than one kind of particle, the criterion of extensivity is ambiguous. For example, suppose we have a gas mixture of distinguishable particles, with N_A particles of type A and N_B particles of type B. The common textbook definition of the entropy as the logarithm of a volume in phase space gives an answer that is not extensive (and incorrect for other reasons[2]). We might try to impose extensivity with the addition of either

$-k_B \ln(N_A!N_B!)$ or $-k_B \ln[(N_A+N_B)!]$. The first choice is the one we want, of course, but the criterion of extensivity does not eliminate the second. If this path were to be taken, at least one more principle would have to be invoked to obtain an unambiguous definition.

Although extensivity is a useful assumption when analyzing the properties of a material, rather than a system with surfaces, it is not essential to either thermodynamics or statistical mechanics, and should not be included as part of the definition of entropy.

XIII. CONSEQUENCES OF THE 12 PRINCIPLES

The principles I have given have led me to the conclusion that Boltzmann's 1877 definition of the entropy as the logarithm of the probability of macroscopic states for composite systems is superior to any other proposed definition. In particular, it is superior to a definition in terms of a volume in phase space that is often found in textbooks for classical statistical mechanics.

If the principles I have presented in this paper are correct, any other valid definition of entropy must turn out to be equivalent to defining entropy in terms of probability.

XIV. GIBBS' PARADOX

Alternatives to Boltzmann's 1877 definition of the entropy have led to problems that have been debated for over a hundred years. The debate has centered on Gibbs' paradox, which refers to a set of old problems in statistical mechanics.[13] The two main problems concern the properties of the entropy of systems of distinguishable particles. In my opinion, they are both easy to resolve on the basis of the principles I have given.

A. Extensivity

The first version of Gibbs' paradox concerns the properties of the entropy as defined in terms of the logarithm of a volume in phase space. Boltzmann's 1877 definition in terms of the logarithm of the probability of a composite system does not have this problem.

If U is the energy, V is the volume, and N is the number of particles, the volume in phase space (often denoted by Ω) consists of all points for which N particles are in a container of volume V with a total energy less than or equal to U. For an ideal gas, this volume is given by

$$\Omega = V^N \frac{\pi^{3N/2}}{\Gamma(3N/2+1)} U^{3N/2}. \tag{1}$$

If the entropy is defined in terms of the logarithm of this volume in phase space,

$$S_\Omega = k \ln \Omega, \tag{2}$$

Stirling's approximation gives an expression for the entropy of the form

$$S_\Omega(U,V,N) = Nk_B \left[\frac{3}{2}\ln\left(\frac{U}{N}\right) + \ln V + \ln X \right], \tag{3}$$

where X is a constant that can be calculated from Eq. (1).

This expression for the entropy, Eq. (3), is not extensive. As explained in Sec. XII, I do not regard the lack of exten-

sivity as a problem in itself. However, Eq. (3) leads to a violation of the second law of thermodynamics.[6] That is a problem!

Consider an ideal gas of N particles in a volume V, and assume that the entropy before inserting the partition is given by Eq. (3). Now insert a partition that divides the system into two equal volumes. The total entropy after inserting the partition is given by twice the entropy of a system half the size of the original one,

$$2S_\Omega(U/2,V/2,N/2)$$
$$= 2\frac{N}{2}k_B \left[\frac{3}{2}\ln\left(\frac{U/2}{N/2}\right) + \ln\left(\frac{V}{2}\right) + \ln X \right]. \tag{4}$$

The change in S_Ω is

$$\Delta S_\Omega = 2S_\Omega(U/2,V/2,N/2) - S_\Omega(U,V,N) = -Nk_B \ln 2. \tag{5}$$

The decrease in entropy predicted by the entropy in Eq. (3) violates the second law of thermodynamics as expressed in the Clausius inequality,[14]

$$\Delta S \geq \int_i^f \frac{đQ}{T}, \tag{6}$$

where i and f refer to the initial and final macroscopic states, before and after inserting the partition. Because $đQ=0$ while the partition is being inserted, the Clausius inequality is violated by Eq. (5). This violation eliminates a definition of the entropy in terms of the logarithm of a volume in phase space from consideration as the entropy of a classical gas.[6]

Boltzmann's definition of entropy in terms of the logarithm of the probability gives exactly the same result for classical particles whether they are distinguishable or not,[2]

$$S_B(U,V,N) = Nk_B \left[\frac{3}{2}\ln\left(\frac{U}{N}\right) + \ln\left(\frac{V}{N}\right) + \ln X \right]. \tag{7}$$

Because Eq. (7) for the entropy does not violate the second law of thermodynamics, there is no paradox and no problem.

B. Continuity

Another problem, which is also known as Gibbs' paradox, concerns the desire for continuity as the interactions between particles in a system go continuously from being measurably different to being the same for all particles.

For example, consider a classical ideal gas with N_A particles of type A and N_B particles of type B. All particles of a given type have the same properties, but these properties are different for type A and type B particles. The entropy of this system differs from the entropy of an ideal gas of $N=N_A+N_B$ particles of a single kind by the amount

$$\Delta S = -k_B \left[N_A \ln\left(\frac{N_A}{N}\right) + N_B \ln\left(\frac{N_B}{N}\right) \right] > 0. \tag{8}$$

Equation (8) is the well-known entropy of mixing.

The concern is that as the differences in the properties of the two types of particles vanish, the entropy of the system changes discontinuously by the entropy of mixing given in Eq. (8).

First of all, it is quite possible for the interactions between particles to be essentially identical, but to still be able to separate them in some way—using differences in diffusion

rates for different isotopes, for example. In that case, the entropy would not change discontinuously as the differences in the interactions vanish.

However, suppose all differences in interactions, masses, etc. could be made to go continuously to zero. At some point, the differences would become smaller than the resolution of our experiments. Nevertheless, at any level of difference in the interactions, we either would or would not be able to measure the difference.

If the entropy were a property of the system (reality)—instead of a description of the system (representation of our knowledge), as argued in Sec. VII—a discontinuity of the entropy would be strange. However, the entropy is given by the probability, which is, in turn, related to our knowledge of the system. There is no problem with our description (or knowledge) of a system changing discontinuously when our information changes discontinuously. If we cannot determine experimentally that there are two different types of particles, then a description that lumps them together will still be correct. Common practice lumps the various isotopes of an element together for most thermodynamic applications. Although different isotopes are clearly distinguishable, the macroscopic predictions are not affected.

The problem of continuity is often expressed in terms of a continuous change from distinguishable to indistinguishable particles. However, such a change is intrinsically discontinuous and does not occur simply because the interactions between the particles become identical.

XV. SUMMARY

I have put forward 12 principles that have led me to conclude that Boltzmann's 1877 definition of the entropy in terms of the logarithm of the probability of macroscopic states of composite systems is superior to all other options.

It would be too much to hope that my arguments will find universal agreement. However, I hope that further discussions will be clarified by an improved understanding of one point of view. Those who might have different points of view have the opportunity to express which of the principles they object to and present their own alternatives.

The issues I have discussed have been the subject of disagreements for well over a century. It might be that, in the end, the conclusions of the scientific community deviate from the principles I have listed here. However, the purpose of this paper will be fulfilled if it paves the way to a final consensus.

[a] Electronic mail: swendsen@cmu.edu

[1] John von Neumann in a conversation with Claude Shannon. Quoted by M. Tribus and E. C. McIrvine "Energy and information (thermodynamics and information theory)," Sci. Am. **225**, 179–188 (1971).

[2] R. H. Swendsen, "Statistical mechanics of classical systems with distinguishable particles," J. Stat. Phys. **107**, 1143–1166 (2002).

[3] R. H. Swendsen, "Statistical mechanics of classical distinguishable particles," *Computer Simulation Studies in Condensed Matter Physics XV*, edited by D. P. Landau, S. P. Lewis, and H. B. Schüttler (Springer-Verlag, Berlin, 2003).

[4] R. H. Swendsen, "Response to Nagle's criticism of my proposed definition of the entropy," J. Stat. Phys. **117**, 1063–1070 (2004).

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

[6] R. H. Swendsen, "Gibbs' paradox and the definition of entropy," Entropy **10**, 15–18 (2008).

[7] R. H. Swendsen, "Footnotes to the history of statistical mechanics: In Boltzmann's words," Physica A **389**, 2898–2901 (2010).

[8] J. F. Nagle, "Regarding the entropy of distinguishable particles," J. Stat. Phys. **117**, 1047–1062 (2004).

[9] See, for example, P. M. Lee, *Bayesian Statistics: An Introduction*, 3rd ed. (John Wiley & Sons, Hoboken, NJ, 2009).

[10] R. H. Swendsen, "Explaining irreversibility," Am. J. Phys. **76**, 643–648 (2008).

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

[12] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (Wiley, New York, 1985).

[13] J. W. Gibbs, *Elementary Principles of Statistical Mechanics* (Yale University Press, New Haven, 1902); *Elementary Principles of Statistical Mechanics* (Dover, New York, 1960) (reprint).

[14] C. M. Van Vliet, *Equilibrium and Non-Equilibrium Statistical Mechanics* (World Scientific, Singapore, 2008), p. 14.

Enhancing the understanding of entropy through computation

Trisha Salagaram[a)]
Department of Physics, University of Pretoria, Pretoria 0001, South Africa

Nithaya Chetty
Department of Physics, University of Pretoria, Pretoria 0001, South Africa and National Institute for Theoretical Physics, Johannesburg 2000, South Africa

We devise an algorithm to enumerate the microstates of a system comprising N independent, distinguishable particles. The algorithm is applicable to a wide class of systems such as harmonic oscillators, free particles, spins, and other models for which there are no analytical solutions, for example, a system with single particle energy spectrum given by $\varepsilon(p,q) = \varepsilon_0(p^2 + q^4)$, where p and q are non-negative integers. Our algorithm enables us to determine the approach to the limit $N \to \infty$ within the microcanonical ensemble, and makes manifest the equivalence with the canonical ensemble. Various thermodynamic quantities as a function of N can be computed using our methods. © *2011 American Association of Physics Teachers.*
[DOI: 10.1119/1.3623416]

I. INTRODUCTION

Entropy is critical to the understanding of statistical physics. However, students have difficulty conceptualizing the meaning of the entropy. Entropy is often presented as a measure of the randomness of a system. Therefore, Monte Carlo methods are an obvious way for studying the entropy.[1–4]

In this paper, we approach the entropy by computing the relevant statistical quantities directly. Our starting point is the enumeration of microstates. This enumeration quickly becomes a challenge because we have to cope with integers that increase exponentially with the system size, and become larger than the largest integer that can be addressed by a computer. A related problem is that the computational time for the obvious brute-force method scales exponentially with the system size, which makes it difficult to study the system for large N.

We devise a hierarchy of algorithms to enumerate the microstates of a system comprising N independent, distinguishable particles. The algorithms are applicable to a wide class of problems such as harmonic oscillators with single particle energy $\varepsilon(p) \propto p$ and non-relativistic particles in one dimension with $\varepsilon(p) \propto p^2$, where p is a non-negative integer. In the microcanonical ensemble our algorithms enable us to determine the approach to the limit $N \to \infty$, and show the equivalence with the canonical ensemble.

We discuss different algorithms, which show students different ways of solving the same problem, with some methods being exponentially more efficient than others. Our presentation encourages students to think carefully about the manner in which the enumeration of microstates can be done, and helps students appreciate how an algorithm, even though mathematically correct, can be impractical.

For small N a plot of the entropy S versus the internal energy E is jagged and thus not differentiable, which makes the temperature T of the system ill-defined because

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V}. \tag{1}$$

We show that a plot of S versus E converges to a smooth, differentiable form for which the temperature is numerically better behaved, which is consistent with the fact that temperature of an isolated system makes sense only in the thermodynamic limit.

The subject of the enumeration of states is well established.[5] Analytical expressions for the number of microstates, $\Omega(N,E)$, with total energy E, exist only for a small number of problems such as a system of N independent harmonic oscillators and a system of N free particles. In this case Stirling's approximation is applicable and enables the derivation of closed form analytical results. For a wide range of other problems, similar analytical solutions do not exist, and numerical methods are needed.

Moore and Schroeder[6] considered a system of harmonic oscillators and noted that the results become cumbersome with E more than about 200 energy units and overflow errors occur if there N is more than a few thousand. Prentis and Zainiev[7] considered systems in the large N limit. In our work, we compute $\Omega(N,E)$ recursively for $N = 1,2,3,\ldots,N_{max}$, where N_{max} can be made sufficiently large to see convergent results. Our algorithm is numerically stable, and the results are achievable in realtime on a desktop computer.[8]

II. THE COUNTING OF MICROSTATES

We consider an isolated system consisting of N identical non-interacting particles with a single-particle energy given by

$$\varepsilon(p, q) = \varepsilon_0(p^2 + q^4), \tag{2}$$

where p and q are non-negative integers. We wish to determine the total number of microstates $\Omega(N,E)$ accessible to the system with total energy E. As far as we know, no analytical solution for $\Omega(N,E)$ is available for this model. Figure 1 shows the single particle spectrum. $\Omega(N,E)$ depends on whether the particles are distinguishable or not, and whether they are bosons or fermions. The model is non-trivial because of the presence of degeneracies, for example, $\varepsilon(1,0) = \varepsilon(0,1) = \varepsilon_0$, and because the energy levels are not evenly spaced, which makes computing $\Omega(N,E)$ nontrivial.

Given $\Omega(N,E)$ the entropy is given by

$$S(N, E) = k_B \ln \Omega(N, E), \tag{3}$$
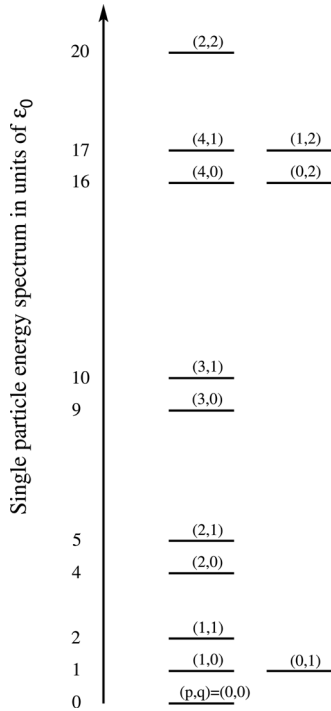
where k_B is the Boltzmann's constant.

Fig. 1. The single particle energy spectrum $\varepsilon(p,q) = \varepsilon_0(p^2 + q^4)$ where p and q are non-negative integers.

A. Pen-on-paper solution

To start with, it is useful to list the system microstates for a small number of particles with a single-particle energy given by Eq. (2). For example, for $N = 3$, and $E = 20\varepsilon_0$, the microstates can easily be listed. Let (p_i, q_i) be the quantum labels for the ith particle, then the system microstates $(p_1,q_1;p_2,q_2;p_3,q_3)$ for indistinguishable spinless particles (bosons) are (2,2;0,0;0,0), (1,2;1,1;0,1), (1,2;1,1;1,0), (4,1;1,1;0,1), (4,1;1,1;1,0), (4,0;2,0;0,0), (4,0;1,1;1,1), (0,2;2,0;0,0), (0,2;1,1;1,1), (3,1;3,1;0,0), (3,1;3,0;0,1), (3,1;3,0;1,0), (3,1;2,1;2,1), (3,0;3,0;1,1), i.e., there are 14 microstates for indistinguishable particles. From these states we can readily infer the corresponding number of microstates for distinguishable particles using the combinatorial factors, for example (2,2;0,0;0,0) $\rightarrow$ 3 complexions for distinguishable particles, (1,2;1,1;0,1) $\rightarrow$ 6 complexions for distinguishable particles, etc. It can therefore be deduced that there are a total of 66 microstates for distinguishable particles, i.e. $\Omega(3,20\varepsilon_0) = 66$. Going through this exercise leaves one with the distinct impression that it is very cumbersome to try to solve this problem using "pen-on-paper" for a large number of particles. An efficient computational algorithm is needed.

B. A brute-force computational method

For distinguishable particles the simplest method involves N nested do-loops, each over the list of single particle energy levels. This method scales exponentially with N. Nevertheless, this method is instructive because it lets students appreciate the rapid increase in $\Omega(N,E)$ with N, and the corresponding exponential increase in computational time. Figure 2 shows a graph of $\ln t$ versus N, where t is the computational time[8] taken to compute $\Omega(N,E)$ for $N = 2,3,4$, and 5 and $E = 500\varepsilon_0$. The graph shows that $t \propto e^{\lambda N}$, where
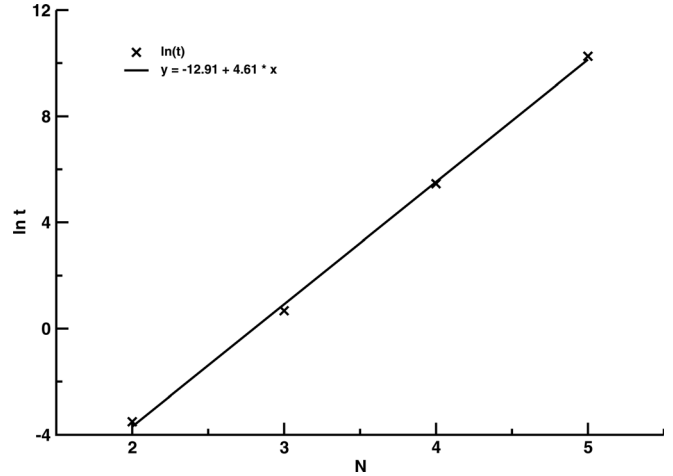

Fig. 2. Graph of $\ln t$ versus N, where t is the computational time in seconds and N is the number of particles. The points are for $N = 2, 3, 4$, and 5.

$\lambda = 4.61$. We conclude that it would take several million years to consider only a moderate number of particles.

C. A recursive algorithm for $\Omega(N,E)$

For fixed total energy E, we consider a system of N distinguishable particles composed of a subsystem of $(N-1)$ particles with $\Omega(N-1, E-E')$ microstates and a subsystem of one particle with $\Omega(1,E')$ microstates (see Fig. 3). In this context $\Omega(1,E')$ is the degeneracy of the single particle levels with energy $E' \leq E$. $\Omega(N,E)$ is constructed by summing the number of ways in which the subsystem of $(N-1)$ particles can be configured with energy $E - E'$ weighted by the degeneracy of the single particle level with energy E', that is,

$$\Omega(N,E) = \sum_{E'=0}^{E} \Omega(N-1, E-E')\Omega(1,E'). \quad (4)$$

Equation (4) is the basis of a recursive algorithm, which can be programmed straightforwardly. $\Omega(1,E')$ can be constructed simply and from $\Omega(1,E')$, $\Omega(N,E)$ may be calculated recursively for $N = 2,3,4,\ldots$.

A plot of the entropy per particle, $s = S/N$, versus $e = E/N$, the energy per particle is presented in Fig. 4 for $N = 3$–10 particles, and compared with the result calculated in the canonical ensemble (see the Appendix) which is in the thermodynamic limit. For these small values of N, the plots are not smooth, which makes the temperature ill-defined.
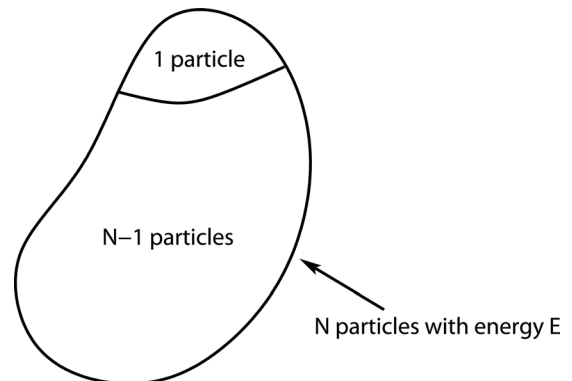
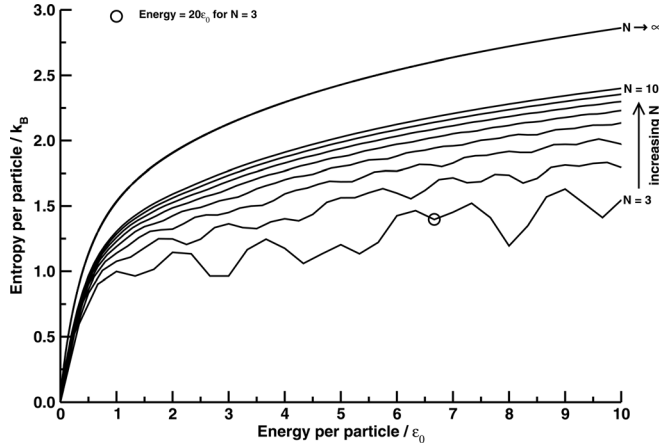

Fig. 3. The system comprising N particles with energy E.

Fig. 4. Entropy per particle versus the energy per particle for $N = 3$–10 distinguishable particles. N increases from bottom to top. The result for $N \to \infty$ is included (top curve) for comparison.

By using Eq. (4), we are able to compute $\Omega(N,E)$ for large values of N. By using a desktop computer,[8] we find that $N_{max} \sim 200$. The computational time is of the order of minutes. For N greater than ~ 200, the largest number addressable by the computer[9] is exceeded, resulting in overflow errors.

Figure 5 shows a plot of s versus e. The result for $N = 200$ has not yet converged to the result for $N \to \infty$. The deviation for s is of order 1% for $e = 10\varepsilon_0$. An improved algorithm is needed.

A small improvement in the algorithm can be achieved by counting states in terms of r, which is the smallest real number that is addressable by the computer. This procedure results in a normalized number of states accessible to the system $\tilde{\Omega}(N,E)$, where $\tilde{\Omega}(N,E) = \Omega(N,E)r^N$. The advantage of counting in terms of r is that a higher value of N is achievable before $\tilde{\Omega}(N,E)$ exceeds the largest number addressable by the computer. By using this method, the maximum number of particles that we were able to consider is $N_{max} \sim 500$. The computational time is of the order of minutes. The result for $N = 490$ in Fig. 5 is improved compared to the value for $N = 200$, but still has not converged compared with the result for the limit $N \to \infty$. The deviations for s are of order 0.5% for $e = 10\varepsilon_0$.
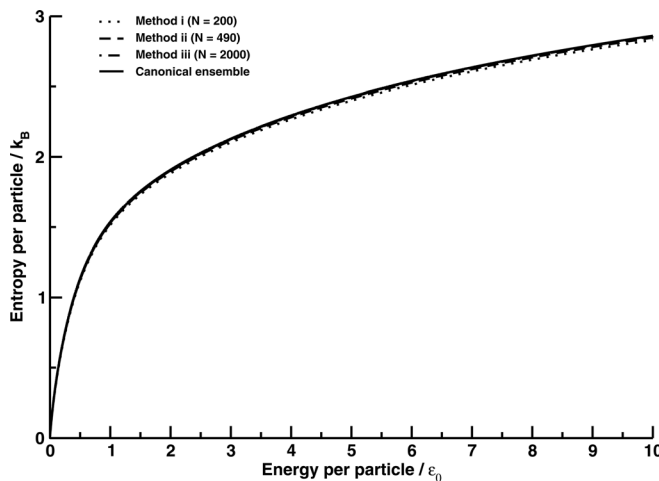

Fig. 5. Entropy per particle s versus e for $N = 200$, 490, and 2000 compared with the result for $N \to \infty$.

D. A recursive algorithm for the entropy S(N,E)

The solution to this problem can be found by realizing that we are interested in the entropy $S(N,E)$ rather than $\Omega(N,E)$ itself. We rearrange Eq. (4) as

$$\Omega(N,E) = \Omega(N-1,E)\left[\Omega(1,0) + \frac{\Omega(N-1,E-1)}{\Omega(N-1,E)}\Omega(1,1)\right.$$
$$\left. + \cdots + \frac{\Omega(N-1,0)}{\Omega(N-1,E)}\Omega(1,E)\right], \quad (5)$$

which gives the following recursive algorithm for the entropy

$$S(N,E) = S(N-1,E) + k_B \ln \sum_{E'=0}^{E} \exp$$
$$\left[\frac{1}{k_B}(S(N-1,E-E') - S(N-1,E))\right]\Omega(1,E'). \quad (6)$$

Working directly with $S(N,E)$ rather $\Omega(N,E)$ is much more manageable computationally. For the range of N and E that we considered, we did not exceed the maximum number addressable by the computer. The computational time for $N = 2000$ and $E = 2 \times 10^4 \varepsilon_0$ is of the order of a day. The result for $N = 2000$ in Fig. 5 is similar to the $N \to \infty$ limit. The deviations for s are less than 0.1% for $e = 10\varepsilon_0$.

The complexity of the problem is illustrated by considering the individual terms in the summation of Eq. (4). In Fig. 6 we have plotted $\Omega(1,E')$ as a function of E'. The curve is discrete because $\Omega(1,E')$ is the degeneracy of the single particle levels with energy E'. We have normalized $\Omega(N-1,E-E')$ by $\Omega(N-1,E)$ for numerical convenience and plotted $\Omega(N-1,E-E')/\Omega(N-1,E)$ and $[\Omega(N-1,E-E')/\Omega(N-1,E)]\Omega(1,E')$ in Fig. 7 as a function of E' for $E = 2 \times 10^4 \varepsilon_0$ and $N = 2000$. The curve of $\Omega(N-1,E-E')$ versus E' is smooth and reflects the convergence of all intensive macrostate variables such as s, the chemical potential μ, and the temperature T.

$\Omega(N,E)$ is constructed by summing over terms of the form of the product of $\Omega(N-1,E-E')$ with $\Omega(1,E')$. These contributions are distinctly discontinuous, and rapidly decreasing as a function of E' as seen in Fig. 7. Figure 7 shows how the microscopic structure of the system underpins its macroscopic properties, even in the limit $N \to \infty$.

III. THE HEAT BATH

The only contributions to $\Omega(N,E)$ come from terms for which $\Omega(1,E') \neq 0$. Therefore, the subsystem of $(N-1)$ particles can only attain energy $E - E'$ if $\Omega(1,E') \neq 0$. This statement holds for any value of N and, in particular, as $N \to \infty$.

However, the main contributions to the summation in Eq. (4) come from terms close to $E' = 0$ for which the subsystem of $(N-1)$ particles has an energy close to E. Hence, all but the low energy single particle states are suppressed. In the large N limit, we may view the system of N particles as being composed of one particle in thermal equilibrium with the $(N-1)$ particle subsystem, and the latter may therefore be viewed as a heat bath. This picture enables us to arrive at the Boltzmann distribution and understand the relation with the canonical ensemble. These observations have important
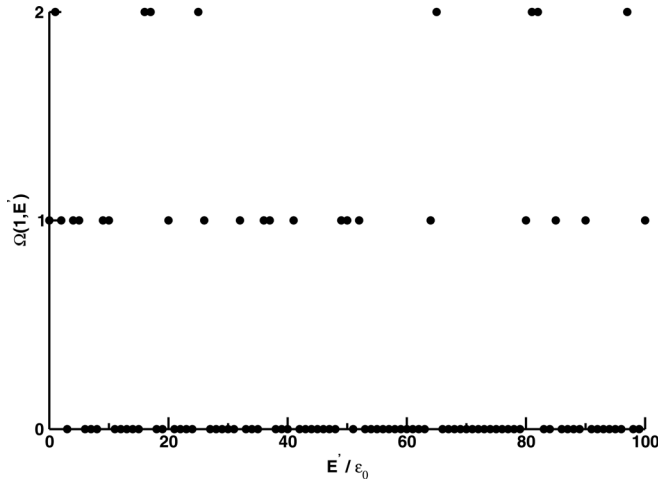
Fig. 6. The number of microstates $\Omega(1,E')$ for $0 \leq E' \leq 100\varepsilon_0$. Its discrete structure means that the terms in Eq. (4) are nonzero only when $\Omega(1,E')$ is nonzero.

implications for the determination of the temperature of the system.

A simple way to determine the temperature as a function of e is to apply Eq. (1) to the curve of $s(e)$. A more interesting way to extract the temperature which shows the equivalence of the microcanonical ensemble with the canonical ensemble is now presented. Following Eq. (6) we have plotted in Fig. 8, the quantity

$$\Delta = \frac{1}{k_B}[S(N-1, E-E') - S(N-1, E)] \qquad (7)$$

for $N = 2000$ as a function of E' for $E = 2 \times 10^4 \varepsilon_0$.

As expected, Δ is negative and linear for small values of E'. The curve becomes more negative for higher values of E', which corresponds to the suppression of the higher energy single particle states. This behavior for higher values of E' is intriguing, and we are not aware that it has been noted before. For small E' we may model this curve by
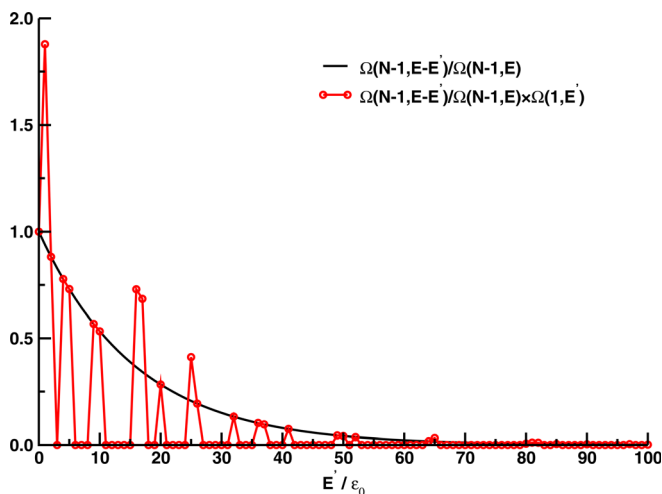
$$\Delta = -\beta E', \qquad (8)$$

Fig. 7. Plot of $\Omega(N-1, E-E')/\Omega(N-1,E)$ versus E' (black curve) and $[\Omega(N-1, E-E')/\Omega(N-1,E)]\Omega(1,E')$ versus E' (circles), where $0 \leq E' \leq 2 \times 10^4 \varepsilon_0$ and $N = 2000$.

so that Eq. (6) may be expressed as

$$S(N, E) = S(N-1, E)$$
$$+ k_B \ln \sum_{E'=0}^{\infty} \exp(-\beta E')\Omega(1, E'). \qquad (9)$$

This form works for the higher energy states as well, which are suppressed even further than what the linear relation in Eq. (8) implies. The upper limit in the summation in Eq. (9) has therefore been replaced by ∞ with no change to the final result. The quantity β is a function of E and emerges here simply from the linear fit to the low energy data in Fig. 8. This observation establishes the equivalence with the canonical ensemble for which β is identified with the inverse temperature $\beta = 1/k_B T$, and $z(T)$ with the single-particle partition function

$$z = \sum_{E'=0}^{\infty} \exp(-\beta E')\Omega(1, E'). \qquad (10)$$

IV. THE CHEMICAL POTENTIAL

Because Eq. (4) corresponds to systematically increasing N one particle at a time, a natural thermodynamic quantity to consider is the chemical potential, which is defined as

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{S,V}. \qquad (11)$$

Because it is difficult to numerically keep the entropy fixed, it is more useful to consider the quantity

$$-\frac{\mu}{T} = \left(\frac{\partial S}{\partial N}\right)_{E,V}. \qquad (12)$$

Because $\delta N = 1$, we may write

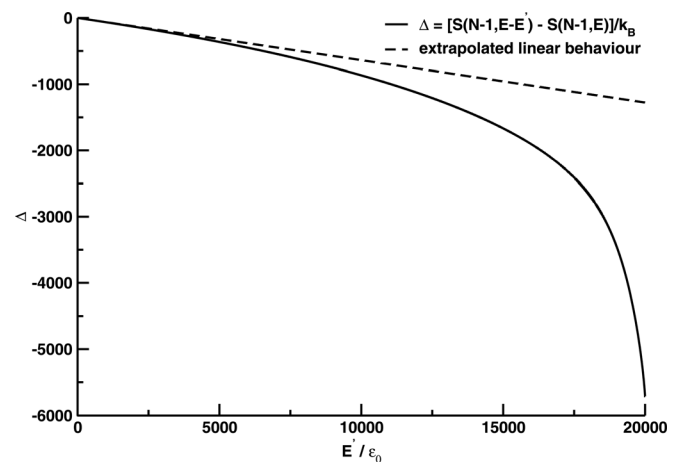$$-\frac{\mu}{T} = S(N, E) - S(N-1, E), \qquad (13)$$

Fig. 8. Plot of $\Delta = (1/k_B)[S(N-1, E-E') - S(N-1, E)]$ as a function of E' for $E = 2 \times 10^4 \varepsilon_0$ and $0 \leq E' \leq 2 \times 10^4 \varepsilon_0$. The curve is linear for low values of E'. The dashed curve is the linear region extrapolated to $E' = 2 \times 10^4 \varepsilon_0$ to show the decrease of Δ for higher values of E', corresponding to the suppression of the higher energy single particle states. The linear region of the curve has a slope equal to $-0.0638\varepsilon_0^{-1}$, which we identify as $-\beta$ according to Eq. (8).
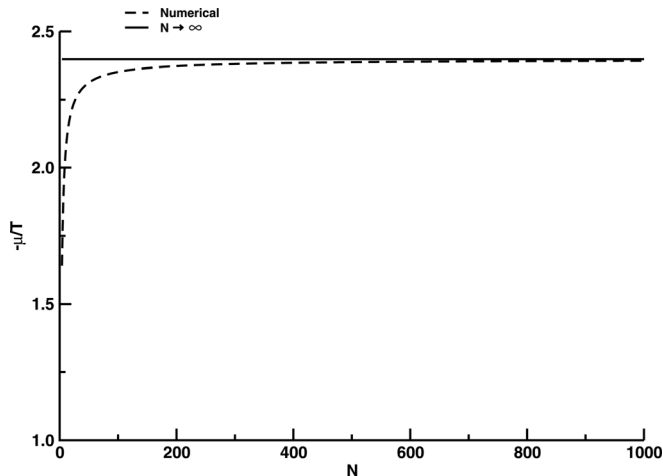
Fig. 9. Plot of $-\mu/T$ as a function of the number of harmonic oscillators, N, at fixed energy per particle equal to $10\varepsilon_0$. The result for $N \to \infty$ limit for μ at the same energy per particle is plotted for comparison. The analytical expression for $N \to \infty$ is $-\mu/T = k_B \ln(e+1)$.

which can be evaluated using Eq. (6). The differentiation in Eq. (12) requires that the total energy E be kept constant. However, if we keep E constant in Eq. (13), $E/N \to 0$ as $N \to \infty$. In this case $s \to 0$ as $N \to \infty$, and Eq. (13) converges to zero. This result is correct but not useful. It is more instructive to keep the energy per particle fixed in Eq. (13), and so that the correct convergence properties of $-\mu/T$ will be found. We leave this problem as an exercise for the reader.

A second subtle point related to the chemical potential in Eq. (13) is that the left-hand side of Eq. (13) is manifestly intensive, whereas the right-hand side appears to be extensive. This can be resolved by investigating the chemical potential using the following scaling argument. Let $s(e)$ be the entropy per particle in the thermodynamic limit. Then

$$S(N,E) = Ns(E/N). \qquad (14)$$

The application of Eq. (12) to Eq. (14) gives

$$-\frac{\mu}{T} = s(e) - e\frac{ds(e)}{de}. \qquad (15)$$

To demonstrate the versatility of our method, we have plotted Eq. (15) for a system of harmonic oscillators as a function of N to test the convergence of $-\mu/T$ (see Fig. 9). The plot is done at constant $e = 10\varepsilon_0$. The result for the chemical potential calculated using Eq. (15) converges to the analytical result for large N. The percentage difference for $N = 2000$ is 0.1%.

APPENDIX: THE CANONICAL ENSEMBLE

We have compared our results for various properties for distinguishable particles in the microcanonical ensemble for finite N with equivalent results in the canonical ensemble

which are, by definition, in the thermodynamic limit. The relevant expressions for the canonical ensemble are given here.

The single-particle partition function $z(\alpha)$ for the model described in Eq. (2) is given by

$$z(\alpha) = \sum_{p=0}^{\infty} \exp(-\alpha p^2) \sum_{q=0}^{\infty} \exp(-\alpha q^4), \qquad (A1)$$

where p and q are non-negative integers, and

$$\alpha = \frac{\varepsilon_0}{k_B T} \qquad (A2)$$

is dimensionless and is a measure of the inverse temperature. $z(\alpha)$ as a function of α can be computed numerically.

The Helmholtz free energy per particle $f(\alpha)$ in units of ε_0 is given by

$$f(\alpha) = -\frac{1}{\alpha}\ln z(\alpha), \qquad (A3)$$

and the average energy per particle $e(\alpha)$ in units of ε_0 is given by

$$e(\alpha) = -\frac{d}{d\alpha}\ln z(\alpha). \qquad (A4)$$

We calculate the average entropy per particle in units of k_B from

$$s(\alpha) = \alpha[e(\alpha) - f(\alpha)]. \qquad (A5)$$

Equation (A5) enables us to plot, for example, the entropy per particle versus the energy per particle at the same value of α.

Another useful result is the expression for the chemical potential in units of ε_0 which is given by $\mu = f(\alpha)$.

a) Electronic mail: trisha.salagaram@up.ac.za

[1] J. R. Drugowich de Felicio and V. L. Libero, "Updating Monte Carlo algorithms," Am. J. Phys. **64**, 1281–1285 (1996).
[2] D. P. Landau, S-H Tsai, and M. Exler, "A new approach to Monte Carlo simulations in statistical physics: Wang-Landau sampling," Am. J. Phys. **72**, 1294–1301 (2004).
[3] J. Tobochnik, H. Gould, and J. Machta, "Understanding temperature and chemical potential using computer simulations," Am. J. Phys. **73**, 708–716 (2004).
[4] J. Tobochnik and H. Gould, "Teaching statistical physics by thinking about models and algorithms," Am. J. Phys. **76**, 353–359 (2008).
[5] Ashley H. Carter, *Classical and Statistical Thermodynamics* (Prentice Hall, Upper Saddle River, NJ, 2001), pp. 251–241; R. K. Pathria, *Statistical Mechanics*, 2nd ed. (Pergamon, Toronto, 1991), pp. 9–26.
[6] T. A. Moore and D. V. Schroeder, "A different approach to introducing statistical mechanics," Am. J. Phys. **65**, 26–36 (1997).
[7] J. J. Prentis and T. Zainiev, "The art of statistical mechanics: Looking at microscopic spectra and seeing macroscopic phenomena," Am. J. Phys. **67**, 1123–1128 (1999).
[8] We used a computer with an Intel Pentium 4 CPU 3.00 GHz and 1 GB of RAM.
[9] For our 32-bit computer, the largest real number addressable by the computer is $1.797693134862316 \times 10^{308}$, and the smallest real number is $2.225073858507201 \times 10^{-308}$.
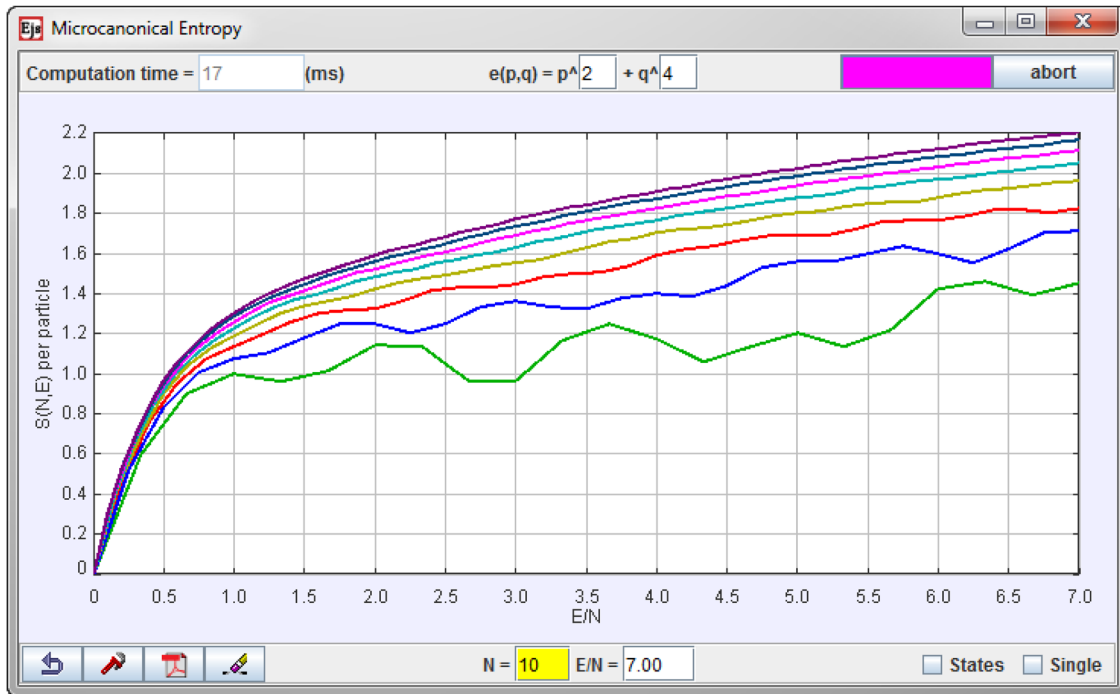
The Microcanonical Entropy model computes the entropy S(N,E) of an isolated system of N identical non-interacting quantum particles with single-particle energy ε given by

$$\epsilon\, p,\, q = \varepsilon_0 \left(p^n + q^n \right)$$

Where the exponents n and m can be varied and units are chosen such that $\varepsilon_0 = 1$. A second window shows the single particle states with energy less than the system energy E.

http://www.compadre.org/osp/items/detail.cfm?ID=11380

The Microcanonical Entropy model is the third in a hierarchy of computational models that demonstrate the computational difficulty of directly calculating the system's microstates $\Omega(N,E)$. The Brute Force Microstates model computes the number of microstates using N nested loops, each over the list of single particle energy levels. This results in a computational scheme that scales exponentially with the system size.

http://www.compadre.org/osp/items/detail.cfm?ID=11353

The Recursive Microstates model computes the number of microstates and the entropy per particle using a recursive algorithm that depends only on $\Omega(N-1,E)$ and $\Omega(1,E)$. This algorithm is fast but the model's long integer arithmetic fails for moderate system size if the number of microstates $\Omega(N,E)$ exceeds the maximum value of the long integer data type 263 -1.

http://www.compadre.org/osp/items/detail.cfm?ID=11381

Wolfgang Christian, Trisha Salagaram, and Nithaya Chetty