The Boltzmann Distribution and Pascal's Triangle

Erich Friedman[†] and W. Tandy Grubbs[‡,*]

Department of Math and Computer Science and Department of Chemistry, Stetson University, DeLand, FL 32720, wgrubbs@stetson.edu

Received October 17, 2002. Accepted December 10, 2002.

Abstract: While Ludwig Boltzmann's contributions to theoretical science are many, the Boltzmann distribution formula is arguably his most important contribution to the field of chemistry. The formula predicts the energy distribution of molecules in an equilibrium system at a given temperature. This distribution in turn governs numerous chemical phenomena, including the intensities of spectral lines, the rates of chemical reactions, and the sedimentation rates of macromolecules, just to name a few. The Boltzmann distribution is normally introduced to students during the statistical mechanics portion of the physical chemistry curriculum. The typical textbook derivation involves applying Lagrange's method of undetermined multipliers to a large system of particles. Undergraduate students, whose mathematics background is often limited to elementary differential and integral calculus, often have trouble grasping this approach because it involves the application of calculus to a multivariable statistical mechanical system. Offered here is an alternate derivation of the Boltzmann distribution formula that instead draws upon some intriguing connections to number theory. The derivation is rigorous and is also pedagogically attractive because one can refer to various graphical energy representations that illustrate how many different ways the particles can be arranged to satisfy the constraints of the system.

Introduction

Ludwig Boltzmann was influential in the development of the kinetic theory of gases and provided much of the foundation of statistical mechanics. He is perhaps most remembered for the absolute entropy formula,

$$S = k\ln\Omega \qquad (1)$$

where k is Boltzmann's constant, 1.381×10^{-23} J·K^{-1}, and Ω is the number of ways of arranging particles in a system to achieve a particular macroscopic state (i.e., the number of *microstates* consistent with a particular *macrostate*). Equation 1 provides a simple statistical definition of entropy, which in turn leads to a statistical interpretation of the second law of thermodynamics; namely, when a constraint is lifted from a closed system, it will evolve toward the most probable macroscopic state; the macroscopic state with the largest value of Ω. The entropy of this new unconstrained state, as defined by eq 1, is guaranteed to be higher than that of the constrained state because the later state has all the original microstates available, plus a new set of microstates due to the absence of the constraint ($\Omega_{\text{unconstrained}} > \Omega_{\text{constrained}}$).

Boltzmann is also credited with deriving the Boltzmann distribution formula, which describes the energy distribution of particles within the maximum entropy state of a system (more commonly referred to as the equilibrium state). The Boltzmann distribution predicts that the population of particles in a particular energy level (ε_j) is proportional to $e^{-\beta\varepsilon_j}$, where β is a constant that is inversely related to the absolute temperature of the system. Others have shown that $\beta = 1/kT$; for example, the quantum evaluation of this constant has been illustrated in reference 1. The Boltzmann distribution has far reaching applications in chemistry. The distribution formula predicts the relative intensities of spectroscopic transition lines (for example, the slight population differences between the spin states of nuclei in a magnetic field that are directly related to the intensity of NMR spectra lines). Other applications of the distribution formula include the field of kinetics (transition state and collision theory) [2], Debye–Hückel theory [3], and the determination of macromolecular weights by sedimentation methods [4].

A number of schemes have been employed to derive and introduce the Boltzmann distribution equation in the undergraduate curriculum. The most common approach for deriving the distribution involves finding the configuration of particles in the system that yields the maximum value of Ω, subject to the constraints of constant particle number (N) and constant internal energy (U) [5–8]. Many undergraduate students have difficulty grasping this approach, especially those students that lack an appropriate background in differential calculus (a background that would allow them to comfortably apply Lagrange's method of undetermined multipliers to a multivariable statistical mechanical system). Other approaches have been proposed to introduce the Boltzmann distribution. One method derives the exponential form of the distribution from eq 1 and the thermodynamic relationship $dS = dU/T$ [9]. Another approach infers the exponential nature of the Boltzmann distribution through the derivation of the barometric formula [10]. Finally, one can infer the exponential form of the distribution formula by examining the relationship $F(\varepsilon_1 + \varepsilon_2) = F(\varepsilon_1)F(\varepsilon_2)$, where $F(\varepsilon_j)$ is the independent probability that a particle in the system has energy ε_j and $F(\varepsilon_1 + \varepsilon_2)$ is the resultant probability that the total energy of the system is equal to $\varepsilon_1 + \varepsilon_2$. The only mathematical

* Address correspondence to this author.
† Department of Math and Computer Science
‡ Department of Chemistry

© 2003 *The Chemical Educator*, S1430-4171(03)02665-3, Published on Web 2/3/2003, 10.1333/s00897030665a, 820116tg.pdf

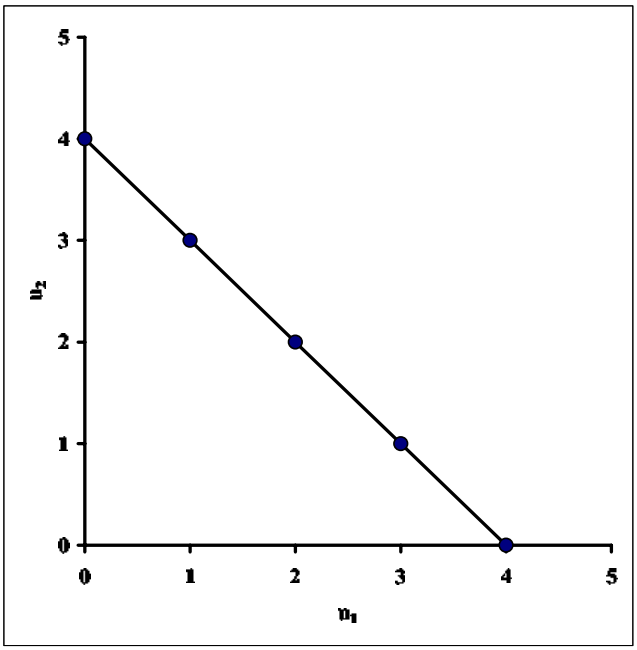

Figure 1. Two-dimensional state-space for a system of two particles ($N = 2$) that can each occupy an energy $u_k = j\varepsilon$ where $j = 0$, 1, 2, 3, or 4 ($m = 4$). The points on the diagonal line illustrate those microstates that satisfy the constraint of constant internal energy ($U = m = 4$).

function that satisfies this probability rule is an exponential function [11].

Offered here is an alternate derivation of the Boltzmann distribution formula that draws upon some intriguing connections to number theory. The derivation is rigorous and the arguments are pedagogically attractive because one can refer to various graphical energy representations that illustrate how many different ways the particles can be arranged to satisfy the constraints of the system.

The System

Consider a macroscopic system consisting of N particles where the thermodynamic internal energy (U) of the system is constant. Furthermore, we will assume that each particle may occupy a series of equally spaced nondegenerate energy levels where the ground state energy level is defined as zero for convenience. The quantum-energy-level spacing (ε) is assumed to be much smaller than the average particle energy ($\varepsilon << U/N$), which guarantees there will be a large number of ways of partitioning the energy in the system to achieve U. While random kinetic motion causes a continuous repartitioning of energy among the particles in the system (sampling all possible microstates with equal probability), the energy of a particular particle k at any instant can be expressed as $u_k = j\varepsilon$ where j is an integer between 0 and m. The constant m is related to the total internal energy of the system as $U = m\varepsilon$. The population of particles in a particular energy state $j\varepsilon$ is represented by n_j.

We now take on the task of determining how many different ways one can arrange the N particles among the $m + 1$ energy levels such that the constraints of constant particle number $\Sigma n_j = N$ and constant internal energy $\Sigma j\varepsilon n_j = U$ are satisfied. The counting of microstates in the system described above is facilitated by representing each permitted arrangement as a point in an N-dimensional "state-space." Beginning with a very simple system of only two particles ($N = 2$), the permitted microstates are represented by the points along the diagonal line in Figure 1 (where m is arbitrarily set equal to 4 for the sake of illustration). Five microstates are available in a system with $N = 2$ and $m = 4$, corresponding to the configurations (4,0), (3,1), (2,2), (1,3), and (0,4).

Our ultimate goal is to derive a probability distribution function for the system, a function that predicts the probability that a particle in the system (we will consider particle number 1 for the sake of argument) resides in the jth energy level. Assuming all microstates have an equal chance of being visited, then the probability distribution function is given by the ratio

$$p(j) = \frac{G(j)}{X} \qquad (2)$$

where $G(j)$ is the number of microstates in which particle number 1 resides in the jth energy level and X is the total number of microstates. In the $N = 2$ system, there is only one microstate associated with particle number 1 having any one of the permitted energies between $j = 0$ and $j = m$. Consequently, the numerator of eq 2 is equal to 1 and the probability distribution function becomes $p(j) = 1/(m + 1)$; the probability distribution is flat.

The microstates available to a three-particle system ($N = 3$) can be represented by the points on the diagonal plane shown in Figures 2A and 2B (where m is again arbitrarily set equal to 4). An examination of Figure 2 reveals a total of 15 microstates ($X = 15$) for the $m = 4$ case. There are $m + 1$ microstates with particle number 1 in the $j = 0$ energy level, $m + 1 - 1$ microstates with particle number 1 in the $j = 1$ energy level, m $+ 1 - 2$ microstates with particle number 1 in the $j = 2$ energy level, and so forth, until one reaches the single microstate where all the system's energy is contained in particle number 1 (the $j = m$ energy level). The probability distribution is not flat for the $N = 3$ system, and instead it decreases linearly with increasing energy level. The probability distribution function for this case, which will be developed formally below, is

$$p(j) = \frac{2(m - j + 1)}{(m + 2)(m + 1)} \qquad (3)$$

Using Pascal's Triangle to Count Microstates

The generation of probability distribution functions for systems containing an arbitrary number of particles will require generic expressions for calculating $G(j)$ and X in eq 2. Mathematicians have long been fascinated with problems of finding the number of ways of arranging N objects among available slots, subject to certain types of constraints. Accordingly, there is a direct connection between the work of early number theorists and the counting of microstates in statistical mechanical systems. Relevant to the system described here is an arrangement of numbers that was intensely investigated by the French mathematician Blaise Pascal over 350 years ago. This pattern of numbers, popularly referred to as Pascal's triangle [12, 13], is shown in Figure 3. Each number in the triangle, $B(r,c)$, is the sum of the number immediately above, $B(r - 1,c)$, and the number immediately above and to the left, $B(r - 1,c - 1)$, where r and c identify the

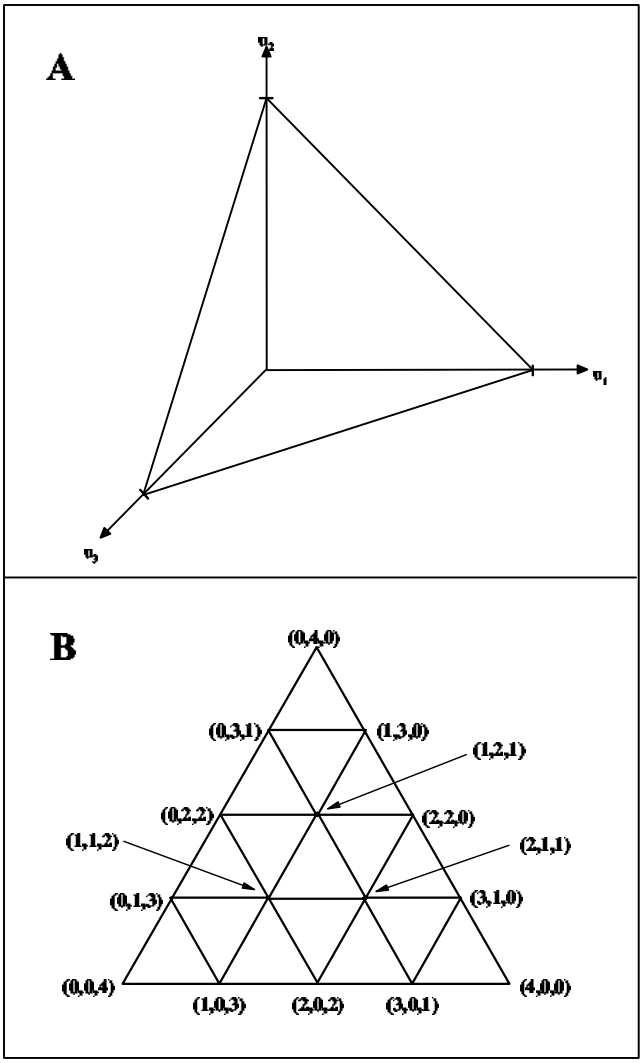

Figure 2. Three-dimensional state-space for a system of three particles ($N = 3$), illustrating the allowed microstates.

location of the number by row and column, respectively. By convention, $r = 0$ and $c = 0$ identify the first row and column of the triangle. The numbers along any row of the triangle yield the binomial coefficients found in the expansion $(x + y)^r$. A useful expression for obtaining any binomial coefficient in the triangle is

$$B(r,c) = \frac{r!}{c!(r-c)!} \qquad (4)$$

where $x!$ denotes the product $x(x-1)(x-2)(x-3)\ldots(1)$ and $0! = 1$ by definition [12].

A remarkable link can be recognized concerning the number of microstates in the systems under consideration here and the binomial coefficients in Pascal's triangle. For the case $N = 3$ and $m = 4$ shown in Figure 2, the total number of microstates (X) is identified by the binomial coefficient $B(6,2) = 15$ (i.e., the binomial coefficient located at $r = 6$ and $c = 2$). Furthermore, the corresponding values of $G(j)$ from eq 2 for the $N = 3$ system are found along the $c = 1$ column of the triangle, yielding $G(0) = B(5,1) = 5$, $G(1) = B(4,1) = 4$, $G(2) = B(3,1) = 3$, and so forth to the top of the column. The values of X and $G(j)$ can be found in Pascal's triangle for $N = 3$ systems described by any value of m; more generally, $X = B(m+2,2)$ and $G(0) = B(m+1,1)$.

We now consider the $N = 4$ system with a goal of verifying that Pascal's triangle can be used to count microstates for a system with any number of particles. Just as one can utilize a two- and a three-dimensional graph to illustrate the microstates of an $N = 2$ and $N = 3$ system, respectively, the microstates for an $N = 4$ system can be represented as points on a four-dimensional hyperplane. (The word hyperplane is used here to refer to any N-dimensional surface that illustrates the available state-space.) The counting of microstates becomes difficult for $N > 3$ systems because the available state-space cannot be easily visualized. Alternatively, the $N = 4$ system can be imagined as a sequence of $N = 3$ triangular planes stacked one on top the other (as shown in Figures 4A and 4B for the case $m = 4$). The point labeled 1 in Figure 4A corresponds to the microstate where all the energy is contained in particle number 1 [14]; the configuration (4,0,0,0). Immediately under this point is a trigonal plane (labeled 2) that illustrates the three configurations (3,1,0,0), (3,0,1,0), and (3,0,0,1); these microstates lie on the three vertices of the triangle. The microstates corresponding to particle number 1 possessing energies $j = 2$, $j = 1$, and $j = 0$ are shown in more detail in Figure 4B, indicating 6, 10, and 15 microstates, respectively. A careful examination of Figures 4A and 4B reveals a total number of microstates equal to 35 and the individual values of $G(j)$ are 15, 10, 6, 3, and 1. Again, we find these numbers in Pascal's triangle; $X = B(7,3) = 35$ and the various values of $G(j)$ are found in the $c = 2$ column.

An important theme begins to emerge as one considers systems with an increasing number of particles. Namely, the microstates for an N-particle system can always be represented as a stacked sequence of $N-1$ hyperplanes. Figures 4A and 4B clearly show how to build the $N = 4$ state-space out of a sequence of $N = 3$ trigonal planes. Another example is provided by the $N = 3$ system; the microstates lying on the trigonal plane of Figure 2 can be imagined as a stacked sequence of $N = 2$ diagonal lines. Crucial to our argument here is the fact that a similar "stacking" relationship is observed in the numbers along any column of Pascal's triangle. For example, the sum of the first five numbers in the $c = 2$ column $1 + 3 + 6 + 10 + 15 = 35$, which is the binomial coefficient found immediately below and to the right of the last number in the sum. This property among the binomial coefficients in the triangle is always true and can be more generally represented as

$$\sum_{r=0}^{m} B(r,c) = B(m+1,c+1) \qquad (5)$$

In summary, we have established a recurrence relationship for generating the N particle state-space out of $N-1$ state-spaces, and we have also shown that this recurrence relationship is reflected in the numbers along adjacent columns of Pascal's triangle, meaning we are free to use eq 4 to calculate the quantities X and $G(j)$ for any combination of N and m. The $N = 3$ and $N = 4$ cases considered above suggest that the total number of microstates can always be represented by the binomial coefficient

		c															
		0	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***	***13***	***14***	***15***
r	***0***	1															
	1	1	1														
	2	1	2	1													
	3	1	3	3	1												
	4	1	4	6	4	1											
	5	1	5	10	10	5	1										
	6	1	6	15	20	15	6	1									
	7	1	7	21	35	35	21	7	1								
	8	1	8	28	56	70	56	28	8	1							
	9	1	9	36	84	126	126	84	36	9	1						
	10	1	10	45	120	210	252	210	120	45	10	1					
	11	1	11	55	165	330	462	462	330	165	55	11	1				
	12	1	12	66	220	495	792	924	792	495	220	66	12	1			
	13	1	13	78	286	715	1287	1716	1716	1287	715	286	78	13	1		
	14	1	14	91	364	1001	2002	3003	3432	3003	2002	1001	364	91	14	1	
	15	1	15	105	455	1365	3003	5005	6435	6435	5005	3003	1365	455	105	15	1

Figure 3. Pascal's triangle. The integers in the triangle are called binomial coefficients and their locations are identified by their row (*r*) and column (*c*) number.

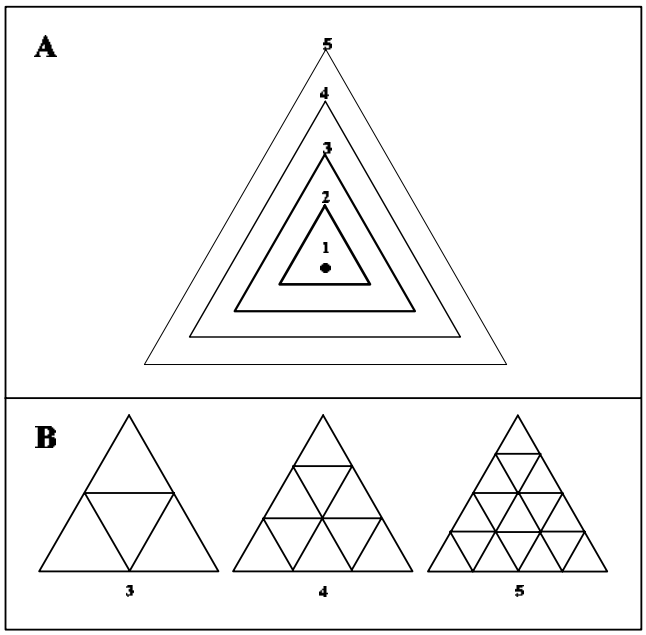

Figure 4. Four-dimensional state-space for a system of four particles (N = 4) that is constructed from a stacked sequence of three-dimensional triangular planes.

$$X = B(N+m-1, N-1) = \frac{(N-1+m)!}{(N-1)!m!} \qquad (6)$$

The corresponding values of $G(j)$ are given by

$$G(j) = B(N+m-2-j, N-2) = \frac{(N-2+m-j)!}{(N-2)!(m-j)!} \qquad (7)$$

The Probability Distribution Function for a Large System

We are now prepared to consider the form of the probability distribution function for a macroscopic system (where N may be on the order of 10^{23}) and ultimately confirm Boltzmann's form for the distribution function that exhibits an exponential energy dependence ($e^{-\beta\varepsilon_j}$). Substituting eqs 6 and 7 into eq 2, one obtains a general expression of the distribution function

$$p(j) = \frac{\left[\dfrac{(N-2+m-j)!}{(N-2)!(m-j)!}\right]}{\left[\dfrac{(N-1+m)!}{(N-1)!m!}\right]} =$$

$$(N-1)\left[\frac{m!}{(m-j)!}\right]\left[\frac{(N-2+m-j)!}{(N-1+m)!}\right] \qquad (8)$$

where the simplification $(N-1)!/(N-2)! = (N-1)$ has been used. Equation 8 can be used at this point to calculate the probability distribution function for the three-particle system considered earlier, confirming eq 3.

Writing out equation 8 for the first few energy levels (j= 0, 1, and 2), one obtains

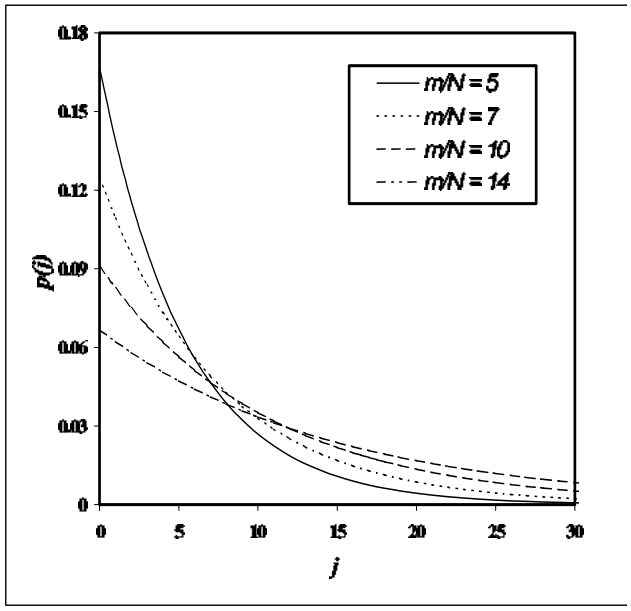

Figure 5. Four plots of the probability distribution function (eq 12) at different temperatures where high temperature is associated with large values of the ratio *m/N*. As expected, the width of the distribution function grows with temperature, implying that the energy of the system is distributed over a larger fraction of particles.

$$p(0) = \frac{(N-1)}{(N-1+m)} \approx \frac{N}{(N+m)} \tag{9a}$$

$$p(1) = \frac{(N-1)m}{(N-1+m)(N-2+m)} \approx \frac{N}{(N+m)}\left(\frac{m}{N+m}\right) \tag{9b}$$

$$p(2) = \frac{(N-1)m(m-1)}{(N-1+m)(N-2+m)(N-3+m)} \approx \frac{N}{(N+m)}\left(\frac{m}{N+m}\right)^2 \tag{9c}$$

where a number of simplifications have been made because we are focusing on systems in which N and m are much greater than 1, meaning expressions like $(N-1+m)(N-2+m)$ and $m(m-1)$ can be approximated as $(N+m)^2$ and m^2, respectively. Examining eqs 9a–c, one can infer a general approximate expression for the distribution function, yielding

$$p(j) \approx \frac{N}{N+m}\left(\frac{m}{N+m}\right)^j = \left(1+\frac{m}{N}\right)^{-1}\left(\frac{m}{N+m}\right)^j \tag{10}$$

This exponential form of the distribution function can be converted to the more familiar exponential form by finding the value of x that satisfies the definition

$$e^x = \left(\frac{m}{N+m}\right)^j$$

Taking the natural logarithm of both sides gives

$$x = j\ln\left(\frac{m}{N+m}\right) = -j\ln\left(\frac{N}{m}+1\right) \tag{11}$$

Using this result, the distribution function can be written in the final form

$$p(j) = \left(1+\frac{m}{N}\right)^{-1} e^{-j\beta} \tag{12}$$

where the constant β is used to abbreviate the logarithmic portion of eq 11

$$\beta = \ln\left(\frac{N}{m}+1\right) \tag{13}$$

Equation 12 confirms Boltzmann's result; the probability distribution function contains an exponential energy dependence. Recognizing that the integer value of j, as defined here, represents the energy that a particle can acquire in units of ε (the energy-level spacing) and recognizing that the term $(1+m/n)^{-1}$ is just a constant for a given system, then eq 12 can be rewritten in the familiar form from statistical mechanics [15]

$$p(j) = \frac{1}{q}e^{-\varepsilon_j\beta} \tag{14}$$

where the constant term, represented by $1/q$, is recognized as the reciprocal of the partition function for the system. Indeed, the standard expression for the partition function from statistical mechanics [16] can be used to independently obtain the constant term in eq 12, yielding

$$q = \sum_{j=0}^{m} e^{-\varepsilon_j\beta} = \sum_{j=0}^{m} e^{-j\beta} = 1+e^{-\beta}+e^{-2\beta}+\ldots+e^{-m\beta}$$
$$= 1+e^{-\beta}+e^{-\beta^2}+\ldots+e^{-\beta^m}$$

$$q \approx \frac{1}{1-e^{-\beta}} = \left(1+\frac{m}{N}\right) \tag{15}$$

where the sum-over-states has been treated as an infinite sum because $m \gg 1$, and where we have substituted for β in eq 13 to achieve the final result.

Others have shown that the constant β is equal to $1/kT$ [1], which, when combined with equation 13, yields a definition for absolute temperature in our system

$$T = \frac{1}{k\ln\left(\frac{N}{m}+1\right)} \tag{16}$$

Keeping in mind that m represents the total energy in our system (in units of ε), then eq 16 predicts that T approaches infinity as m approaches infinity and T approaches zero as m approaches zero, as expected. An inspection of eq 16 reveals that the absolute temperature of a system depends on the ratio

m/N; high temperature implies high internal energy (large m), which is associated with a large value of m/N. Figure 5 shows several plots of the probability distribution function (eq 12) at different m/N ratios (temperatures), revealing that the width of the distribution function grows with temperature.

Conclusion

The results obtained so far illustrate that the probability distribution function is indeed exponential in nature for the specific system considered here, a system of particles that may occupy a series of equally spaced nondegenerate energy levels (such as a system of independent harmonic oscillators). Will a similar result be obtained when the energy-level spacing is not constant or when degeneracies are present? The quantum mechanical energy expression for a particle of mass μ that is free to translate in a one-dimensional box of length L is given by

$$E_n = \frac{n^2 h^2}{8\mu L^2}$$

where the quantum number n can take on values 1, 2, 3, ... [17]. Redefining the translational energies relative to the ground-state energy (setting the ground-state energy to zero), the allowed energy levels become

$$\varepsilon_n = (n^2 - 1)\varepsilon \qquad (17)$$

with

$$\varepsilon = \frac{h^2}{8\mu L^2}$$

Equation 17 reveals that the energy levels available to the system are all a multiple of a constant ε, and the allowed multiples are defined by $(n^2 - 1)$, which takes on integer values 0, 3, 8, 15, ..., m for n quantum number values 1, 2, 3, 4, ..., $(m + 1)^{1/2}$, respectively. When compared to our previous treatment, the microstates available in a system of particles that can occupy these translational energy levels can be carried out using the same Pascal's triangle-based method, only the values of j that are considered must be restricted to those satisfying the condition $(n^2 - 1)$; $j = 0, 3, 8, 15, ..., m$. Although the total number of microstates available to the system will be greatly reduced in comparison to the values predicted by eq 6, the use of eq 7 is still valid and a full derivation of the probability distribution function (not shown here) again reveals an exponential dependence on j. Only the constant term in eq 12, representing the partition function, and the magnitude of β are changed by the reduction in total microstates available in the system.

Similar arguments can be made concerning the rotational energy states of a system, where the allowed energy levels can be represented as a constant times the term $J(J + 1)$ where $J = 0, 1, 2, ...$ is the rotational quantum number [18]. The probability distribution function can be inferred for this system by only considering values of j that satisfy the condition $J(J + 1)$, j=0, 2, 6, ... The degeneracy of the rotational levels can be accounted for in the derivation by incorporating the degeneracy factor $(2J + 1)$ in eq 7. This will account for the fact that, for example, the microstates associated with a triply degenerate energy level would need to be counted three times.

The energy levels available to any imaginable system can be described as we have done here by selecting a suitably small energy-level spacing (ε), thereby guaranteeing that the real energy levels will always nearly coincide with one of the jth energy levels. Although the value of m may be unusually large, one can still utilize the approach demonstrated in this work to derive the probability distribution function, which will always reveal an exponential dependence on j (the energy of the level).

References and Notes

1. Hakala, R. W. *J. Chem. Educ.* **1962,** *39,* 525.
2. House, J. E. *Principles of Chemical Kinetics;* Wm. C. Brown Publishers: Boston, MA, 1997; pp 99–103.
3. Barrow, G. M. *Physical Chemistry,* 6th ed.; McGraw-Hill: New York, 1996; pp 420–421.
4. Atkins, P. *Physical Chemistry,* 6th ed.; W. H. Freeman: New York, 1998; pp 686–687.
5. Atkins, P. *Physical Chemistry,* 6th ed.; W. H. Freeman: New York, 1998; pp 571–573.
6. Winn, J. S. *Physical Chemistry;* HarperCollins College Publishers: New York, 1995; pp 829–831.
7. Alberty, R. A.; Silbey, R. J. *Physical Chemistry,* 2nd ed.; Wiley & Sons.: New York, 1997; pp 555–557.
8. Hakala, R. W. *J. Chem. Educ.* **1961,** *38,* 33–55.
9. McDowell, S. A. C. *J. Chem. Educ.* **1999,** *76,* 1393–1394.
10. Levine, I. N. *Physical Chemistry,* 4th ed.; McGraw-Hill: New York, 1995; p 444.
11. Widom, B. *Statistical Mechanics;* Cambridge University Press: Cambridge, UK, 2002: p 5.
12. Ogilvy, C. S.; Anderson, J. T. *Excursions in Number Theory;* Dover Publications: New York, 1966; pp 48–50.
13. Conway, J. H.; Guy, R. K. *The Book of Numbers;* Springer-Verlag: New York, 1996; pp 66–71.
14. The possibility that all the systems energy resides in a single particle becomes exceedingly small as one considers systems with a large number of particles. Indeed, this possibility would be disallowed by relativistic arguments for a macroscopic system ($N \approx 10^{23}$) because it would necessitate a particle having a translational kinetic velocity in excess of the speed of light.
15. Winn, J. S. *Physical Chemistry;* HarperCollins College Publishers: New York, 1995; p 831.
16. Winn, J. S. *Physical Chemistry;* HarperCollins College Publishers: New York, 1995; pp 830–831.
17. Winn, J. S. *Physical Chemistry;* HarperCollins College Publishers: New York, 1995; pp 854–857.
18. Winn, J. S. *Physical Chemistry;* HarperCollins College Publishers: New York, 1995; pp 859–861.

© 2003 *The Chemical Educator*, S1430-4171(03)02665-3, Published on Web 2/3/2003, 10.1333/s00897030665a, 820116tg.pdf