54
8-8

LA-4868

2228

MASTER

A Derivation of the
Boltzmann Distribution Law
with Illustrations from a
Hard-Sphere Model Computer Program

los alamos
scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87544

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Atomic Energy Commission, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

Printed in the United States of America. Available from
National Technical Information Service
U. S. Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22151
Price: Printed Copy \$3.00; Microfiche \$0.95

The logo features a vertical line with a wavy section, ending in a circle containing a smaller circle with a cross inside. Below this is the text "los alamos scientific laboratory of the University of California LOS ALAMOS, NEW MEXICO 87544".

los alamos
scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87544

LA-4868

UC-4 and UC-34

ISSUED: June 1972

A Derivation of the Boltzmann Distribution Law with Illustrations from a Hard-Sphere Model Computer Program

by

Donald R. Conant

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Atomic Energy Commission, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

A handwritten signature, possibly "Jeg", is written over the bottom right corner of the page.

A DERIVATION OF THE BOLTZMANN DISTRIBUTION LAW
WITH ILLUSTRATIONS FROM A HARD-SPHERE MODEL COMPUTER PROGRAM

by

Donald R. Conant

ABSTRACT

The Boltzmann distribution law is derived, and examples illustrating differences among the assumptions of Boltzmann, Bose-Einstein, and Fermi-Dirac statistics are given. The range of applicability of the assumptions and the distortions of molecular systems caused by this limited range are discussed. Results from a hard-sphere model calculated by computer are used to illustrate the distortions and the rather good agreement with the Boltzmann prediction for energies near the average molecular energy even for systems containing only a few molecules.

I. INTRODUCTION

The Boltzmann distribution and, to a lesser extent, the distributions resulting from Bose-Einstein or Fermi-Dirac statistics are discussed in many textbooks and papers (e. g., Refs. 1-12). These derivations, however, are either so brief or so abstract that much of the derivation and, especially, its fundamental assumptions are not thoroughly grasped by students.

The purpose of this report is sevenfold:

(1) to give a better understanding of the statistical assumptions that are the basis for the derivation;

(2) to illustrate the basic differences among Boltzmann, Bose-Einstein, and Fermi-Dirac statistics;

(3) to indicate more completely why the method of Lagrangian multipliers works;

(4) to indicate that certain approximations used in its derivation are responsible for differences between the Boltzmann distribution and the actual distributions for real finite systems, but that these differences become extremely small over the energy range of greatest interest for systems containing many particles;

(5) with the aid of the hard-sphere model,

whose behavior was examined by computer, to illustrate the nature of the deviations of real distributions from the Boltzmann distribution;

(6) to illustrate with the hard-sphere model that the Boltzmann distribution is in fair agreement with a system containing as few as 8 molecules, and in excellent agreement with a system containing 108 molecules;

(7) to explain why the Boltzmann distribution is such a good approximation for systems containing relatively few molecules, whereas the usual derivation requires many molecules to justify the approximations used.

Matter is composed of an enormous number of small particles that constantly move about and interact with each other. Very little is known about the exact character of their motion and interaction. The large number of particles involved makes a direct exact solution of the physical problem quite impossible. Nevertheless, experience has shown that aggregates of many interacting particles are subject to certain laws that arise specifically because of that very large number which makes direct solutions impossible. The discipline "Statistical Mechanics" studies the bulk properties associated with large numbers of particles and predicts

relationships among them on the basis of certain assumptions about probability. Such an approach neglects most of the detailed relationships among the particles and concentrates upon assumptions about probability for the information needed to predict the overall behavior of a system.

The word "probability" can have any of several meanings. Probability can be regarded as:

- (1) the relative frequency of occurrence of an event in a large number of trials;
- (2) a result that can be deduced from the more primitive notion of equal likelihood;
- (3) the degree of certainty of the observer.

In the physical sciences, we are concerned primarily with the first two concepts of probability. Statistical mechanics assigns an equal likelihood to those outcomes that, to the best of our knowledge, are symmetrical and have no special properties that distinguish among them. The validity of such assignments, however, rests with experimental test. Thus, the basis for statistical reasoning involves the second concept of probability, whereas its justification involves the first concept.

The advent of large digital computers provided the means whereby the exact equations governing the motions of a moderate number of particles can be specified and the resulting motions studied. The relationships among bulk properties for such exactly determined systems can be observed and compared with those predicted by statistical mechanics. Such examples illustrate the correctness of assignments of equal likelihood for the models studied. To the extent that the models abstract the important characteristics of real systems, the justification of the method is extended.

II. PROBABILITY

This concept is useful only for systems in which after many trials a trend for any given result to occur with a certain approximate frequency becomes evident. The uncertainty in the value of this frequency should decrease with an increased number of trials. Because of the possibility that a very rare occurrence may predominate even after many trials, statistics does not require absolutely that a large number of trials will lead to a fairly accurate frequency determination. However, statistical theory is only useful for such cases.

Even the concept of equal likelihood is based upon results involving a large number of trials. "Equal likelihood" of two different events means that, in theory, both would occur with the same frequency if it were possible to carry out an infinite number of trials; or, in practice, that both would occur with nearly the same frequency after many trials, neglecting extremely rare events. Again, if the frequency of a common event cannot be determined because of the unfortunate repeated occurrence of a rare event, the usefulness of the statistical approach is lost.

Because statistical mechanics starts by assigning equal likelihood to various events, it is important to mention problems that may arise in using such an approach. Because of the present impossibility of measuring particle-particle interactions and motions directly when many interacting particles are present, one cannot be certain that all important properties distinguishing among the various possible events have been taken into account. Even when all of the properties for a given problem are known, distinguishing characteristics that must be taken into account may be subtle, as is illustrated by Bertrand's paradox:

If a chord of a circle is chosen at random, what is the probability that its length exceeds the length of a side of the inscribed equilateral triangle?

(1) The position of the chord may be fixed by giving its midpoint. Those chords that meet the above conditions have midpoints lying within the concentric circle whose radius is half that of the original circle. Because the area of this circle is one quarter of the area of the original circle, the probability is $1/4$.

(2) The direction of the chord can be fixed in advance because of the symmetry of the problem. If the diameter perpendicular to this direction is drawn, only chords intersecting that diameter in the interval from one quarter to three quarters of its length will satisfy the original conditions. Therefore, the probability is $1/2$.

(3) One end of the chord may be fixed in advance because of the symmetry of the problem. Because the two sides of the inscribed equilateral triangle with its vertex at this point and the tangent to the circle at this point form three 60°

angles, only chords falling within the middle angle satisfy the original conditions, giving a probability of $1/3$.

The three different probabilities result from three different assignments of equal likelihood, each of which has equal justification. The probability that should be used depends upon the method used for randomly choosing chords. Statistical mechanics faces the same problem in trying to decide how, in actual fact, nature chooses randomly among different events.

The equal likelihood assignment applied to physical systems is that of equal a priori probability for equal extensions in phase space. This implies that momentum and volume are fundamental to probability assignments concerning nature, rather than energy, distance, or area. A further implication is that the results are also assumed to be independent of the direction of time.

A further problem involving the assignment of appropriate probabilities to various states of a system is that, theoretically, the final result does depend upon the exact initial condition of the system. The reason that the ergodic hypothesis cannot possibly be obeyed in its original strict form is associated with this fact. Strictly, it states that the phase point for any isolated system would pass in succession through every point compatible with the energy of the system before finally returning to its original position in the phase space.⁵ A simple example that would violate this hypothesis would be gas particles contained in an ideal container and having velocities and initial positions such that they never collided with one another. If collisions with the container wall did not transfer energy to the particles, any arbitrary energy distribution could be imposed upon the particles and would remain stable. The system would not pass through states involving other energy distributions, and for such a system the Boltzmann distribution would be no more likely than any other. Statistical mechanics is useful only when the real situation does not involve a peculiar configuration that prevents passage of the system's phase point through at least every region compatible with its energy. Although the example just mentioned is highly artificial, the exact initial conditions of

a real system can conceivably be important in determining its resultant bulk behavior.

III. BOLTZMANN DISTRIBUTION

The probability assignments normally used in deriving the Boltzmann distribution do not agree with those suggested by quantum theory. This distribution is therefore considered to be only a good approximation for real systems. When absolute entropies are calculated, the Boltzmann theory must be modified to account for the different probability assignments of quantum theory. A state may be described by indicating the quantum states available and the number of particles contained in each. It is customary to segregate the N particles available into groups of N_i particles each of which has the same energy, ϵ_i . All states of nearly the same energy are similarly grouped together.

Because Boltzmann statistics is based upon probability assumptions that correctly apply to certain situations in the macroscopic world, we present an example illustrating these assumptions and the differences between them and quantum mechanical assumptions. Consider a box divided into four compartments of equal size. These four compartments might correspond to four states, three, perhaps, with one energy, ϵ_1 , and the fourth with energy ϵ_2 . If two identical marbles were thrown, one after the other at the box from a long distance and the process were repeated until both marbles had lodged within the box, then no given compartment should be strongly favored over any other. If the number of marbles in each compartment is recorded and the process is repeated, eventually a frequency for each possible distribution of the two marbles in the box will become apparent within certain limits of uncertainty. If the frequency of one particular box's containing two marbles is divided by the frequency of that same box's containing only one marble, the ratio will be approximately $1/6$, which is a prediction that agrees with Boltzmann statistics. If we were talking about molecules and quantum states, however, the prediction would be different. Bose-Einstein statistics, which applies to particles with zero or integral spin, would predict a ratio of $1/3$; Fermi-Dirac statistics, which applies to particles with half-integral spin, would predict a ratio of

0/3. Table I lists the various distributions considered by each statistical system for our example.

Normally particles are considered to be "distinguishable" in the Boltzmann system and "indistinguishable" in the Bose-Einstein or Fermi-Dirac systems. Such terminology is somewhat misleading, because it really makes no difference to the final statistical results determined by trial through throwing marbles at a box whether the observer does not or cannot determine which marble is which. The observer's ability to distinguish between particles is not the fundamental issue. The situations assumed to have equal a priori probability of occurrence are of prime importance. In Boltzmann statistics, a distinction is made among each of the particles in making the equal probability assignments. In the other systems, no such distinction is made. Boltzmann statistics assumes that each marble has equal probability of entering any of the four compartments regardless of where the other marble is thrown. This leads to $4^2 = 16$ ways in which the marbles might be distributed after any trial, and each way is assumed to have an equal probability of occurrence. If, however, the boxes were so small that each could only hold one marble, all ways assuming two marbles in one box would have to be eliminated. Other changes could require still different statistical assumptions. Thus, in trying to predict behavior of quantum systems, the probability distribution of one particle is not assumed to be entirely independent of the other particles present.

We shall now consider how the number of equally likely distributions may be calculated by the three statistical methods for a system with N particles and g states. For the Boltzmann method, each particle can be placed in any of the g states with equal likelihood. For each of the g ways in which the first particle is placed, the second particle may be placed in g ways, giving $g \cdot g$ or g^2 ways in which two particles can be distributed. For three particles, we have $g \cdot g \cdot g = g^3$; and for N particles, there are g^N ways. Thus, in our example there are $4^2 = 16$ ways.

The Bose-Einstein method makes no distinction among the particles, and equal probability is assigned to each distribution possible under these circumstances. If s_i represents state i and p_i

represents particle i , the first distribution can be represented as

$$s_1 p_1 p_2 s_2 s_3 s_4 ,$$

where particles to the right of any state are considered to be in that state. Leaving s_1 always at the left, there are for the general case $(N + g - 1)!$ ways in which the remaining $N + g - 1$ symbols can be arranged, because the second symbol can be placed in $N + g - 1$ positions, the next in $(N + g - 1) - 1$ positions, the fourth in $(N + g - 1) - 2$, etc., down to the last which has only one remaining position. Thus,

$$(N + g - 1)(N + g - 2)(N + g - 3) \dots (2) \cdot (1) \\ = (N + g - 1)!$$

Because, in the final analysis, we do not wish to distinguish among the particles, those rearrangements that only exchange p_i 's must be eliminated. Because there are N particles, $N!$ ways must be eliminated by dividing $(N + g - 1)!$ by $N!$

Furthermore, any rearrangement that exchanges an s_i together with all of its p 's, with s_j and its p 's, does not constitute a new arrangement of particles among the states. For example,

$$s_1 p_1 p_2 s_2 p_3 s_3 p_4 s_4 ,$$

and

$$s_1 p_1 p_2 s_4 s_3 p_4 s_2 p_3 ,$$

where $s_2 p_3$ has been exchanged with s_4 , both represent particle 3 in state 2 and no particle in state 4. Because there are $g - 1$ states that can be exchanged in this manner, $(g - 1)!$ must also be divided into the previous result to eliminate these duplications. The total number of ways assigned equal probability for Bose-Einstein systems is then $(N + g - 1)! / N!(g - 1)!$. Thus, in our example there are

$$\frac{(2 + 4 - 1)!}{2!(4 - 1)!} = \frac{5!}{2!3!} = 10 \text{ ways} .$$

TABLE I

DISTRIBUTIONS FOR A SYSTEM OF TWO PARTICLES AND FOUR STATES

Boltzmann					Bose-Einstein					Fermi-Dirac				
Ways of Distribution	States				Ways of Distribution	States				Ways of Distribution	States			
	ϵ_1		ϵ_2			ϵ_1		ϵ_2			ϵ_1		ϵ_2	
	1	2	3	4		1	2	3	4		1	2	3	4
1	AB				1	••								
2	A	B			2	•	•			1	•	•		
3	A		B		3	•		•		2	•		•	
4	A			B	4	•			•	3	•		•	
5	B	A												
6		AB			5		••							
7		A	B		6		•	•		4		•	•	
8		A		B	7		•		•	5		•	•	
9	B		A											
10		B	A											
11			AB		8			••						
12			A	B	9			•	•	6			•	
13	B			A										
14		B		A										
15			B	A										
16			AB		10				••					

The Fermi-Dirac method makes no distinction among the particles, and no more than one particle is allowed in any given state. N of the g states will be occupied. If the symbols representing the states were placed in a row, we may require that the first N states be occupied, the remaining states being empty. Thus,

$$s_5 s_1 s_3 s_{257} \text{ ---- } s_8 \quad s_{22} \text{ ---- } s_{197}$$

$$1 \ 2 \ 3 \ 4 \ \text{----} \ N \quad N+1 \ \text{----} \ g$$

There are $g!$ ways to rearrange the s 's; however, rearrangements that exchange only the first N states among themselves, or the last $g-N$ states among themselves, do not yield a new distribution when no distinction is made among particles. The total number of ways assigned equal probability for Fermi-Dirac systems is then $g! / N!(g-N)!$. In our example there are

$$\frac{4!}{2!(4-2)!} = \frac{4!}{2!2!} = 6 \text{ ways}$$

We wish to find the most probable distribution of particles among the various energy levels. Therefore, we divide the N particles into groups of N_i particles, each of which has the same energy, ϵ_i . The number of ways in which any given distribution of this sort can be obtained is required. In the Bose-Einstein and Fermi-Dirac methods, this number is simply the product of the number of ways for each energy level. Thus, if W_E represents the first and W_F the second, then

$$W_E = \prod_i \frac{(N_i + g_i - 1)!}{N_i! (g_i - 1)!},$$

and

$$W_F = \prod_i \frac{g_i!}{N_i! (g_i - N_i)!}.$$

In our example if we were interested in the number of ways in which the distribution could be obtained with two particles in ϵ_1 and no particles in ϵ_2 , the results would be $N_1 = 2$, $g_1 = 3$, $N_2 = 0$, $g_2 = 1$,

$$W_E = \frac{(2+3-1)!}{2!(3-1)!} \cdot \frac{(0+1-1)!}{0!(1-1)!}$$

$$= \frac{4!}{2!2!} \cdot \frac{0!}{0!0!}$$

$$= 6$$

because $0! = 1$, and

$$W_F = \frac{3!}{2!(3-2)!} \cdot \frac{1!}{0!(1-0)!}$$

$$= \frac{3!}{2!1!} \cdot \frac{1!}{0!1!}$$

$$= 3$$

In the Boltzmann system, although the factor $g_i^{N_i}$ represents the number of ways in which a distribution can be made up from N_i specified molecules in energy level ϵ_i , it does not take into account the additional ways that can be obtained by exchanging particles among energy levels. The number of ways in which N particles can be exchanged is $N!$. But exchanges within any given energy level have already been taken into account by the factors $g_i^{N_i}$; therefore, the term $N!$ must be divided by each $N_i!$ to prevent counting these exchanges a second time. The number of ways in which particles can be exchanged among energy levels for any given energy distribution is $N! / \prod_i N_i!$, and we have

$$W_B = N! \prod_i \frac{g_i^{N_i}}{N_i!}$$

for the number of ways, W_B , in which any given energy distribution can be obtained by the Boltzmann method. For the energy distribution with $N_1 = 2$, $g_1 = 3$, $N_2 = 0$, $g_2 = 1$, then

$$W_B = 2! \left(\frac{3^2}{2!} \cdot \frac{1^0}{0!} \right)$$

$$= 9$$

Our example distribution did not involve exchange among energy levels because all particles were contained in ϵ_1 . The energy distribution with $N_1 = 1$, $g_1 = 3$, $N_2 = 1$, $g_2 = 1$ does allow for such exchanges, giving

$$W_B = 2! \left(\frac{3^1}{1!} \cdot \frac{1^1}{1!} \right)$$

$$= 6$$

If $g_i \gg N_i$, then the vast majority of distributions have only one particle contained in each state and those having more than one particle in a state can be neglected. Neglecting exchanges among energy levels, Boltzmann statistics gives $\prod_i g_i^{N_i}$ for the number of ways in which an energy distribution can be achieved. However, this includes exchanges among particles within each energy level. These may be eliminated approximately by dividing by $N_i!$. This correction is not exact because some of the states present actually contain more than one particle in some distributions, and true Boltzmann counting never assumes that exchanging particles in the same state leads to a different equally likely way of arriving at a distribution. "Corrected" Boltzmann counting is thus given by

$$W_B^{\text{Corr.}} = \prod_i \frac{g_i^{N_i}}{N_i!}$$

and is a good approximation for W_E only when $g_i \gg N_i$. Under these conditions, W_E and $W_F \approx W_B^{\text{Corr.}}$ because none of the three "distinguishes" among particles, and the only differences among them arise because of differences in treating those distributions that contain more than one particle per state.

In the example $N_1 = 2$, $g_1 = 3$, $N_2 = 0$, $g_2 = 1$, corrected Boltzmann counting gives

$$W_B^{\text{Corr.}} = \frac{3^2}{2!} \cdot \frac{1^0}{0!} = \frac{9}{2} = 4.5$$

If there is no distinction among particles, the Bose-Einstein value of $W_E = 6$ should be correct. The error in $W_B^{\text{Corr.}}$ is caused by the fact that the three distributions with both particles in either state 1, state 2, or state 3 are divided by 2 as though in the original Boltzmann counting the distribution AB, —, —, — was counted as different from the distribution BA, —, —, —, etc., which is not the case. Adding 3/2 back in gives 4.5 + 1.5, or 6, the correct result.

The most probable energy distribution is that which can be formed by the largest number of equally likely ways. In our example, there are three different possible energy distributions for a Boltzmann system: $N_1 = 2, N_2 = 0$; $N_1 = 1, N_2 = 1$; or $N_1 = 0, N_2 = 2$. The number of ways in which each of these can be made up is, respectively: 9, 6, and 1. For this example, then, the most probable energy distribution is $N_1 = 2, N_2 = 0$ because it can be made up by the largest number of ways, 9.

If many particles and an infinite number of energy levels are involved, one must turn to the methods of calculus to maximize the function W_B . Because factorials are difficult to use with this method, Stirling's approximation for $\ln N!$ is usually introduced to substitute a more tractable function. As shown in Fig. 1, the areas under the two curves bracket the value for the area under the discontinuous stair-step line. Because the area under the stair-step function equals $\ln N!$,

$$\int_0^N \ln(x+1) dx > \ln N! > \int_1^N \ln x dx,$$

or

$$(N+1)\ln(N+1) - N > \ln N! > N \ln N - N + 1.$$

If $N \gg 1$, unity can be neglected, and the upper and lower limits approach the same value of $N \ln N - N$. So for large N , we can use the relationship

$$\ln N! \approx N \ln N - N$$

In a real system, only a finite number of particles is involved. Because the distribution law to be derived for the translational energy contains an exponential term, all values of N_i for a real finite system will not be large compared to unity for translational energy states, as is necessary to justify the use of Stirling's approximation. Very high energy states will be nearly or completely empty.

It is because of the energy restriction that most of the molecules are required to be in the lower energy levels. When each ϵ_i covers a range such that the lower levels contain vast numbers of particles of nearly the same energy, the major contribution to $\ln W_B$ is from these levels. The greatest error comes from levels containing from one to only a few particles, and it is small if there are many particles in the system and the energy range for each ϵ_i is considered large enough, as previously mentioned. The empty levels do not contribute any error to $\ln W_B$ when Stirling's approximation is used. Stirling's approximation is exact for empty energy levels because

$$\ln 0! = 0 = \lim_{x \rightarrow 0} x \ln x - x$$

The expectation of a distribution derived using an approximation for $\ln N_i!$ that is good only for large N_i , is that it will fairly well approximate the true distribution for systems of interest in the energy ranges with moderate to high particle populations.

Fig. 1. Plots of a stair-step function, $\ln(x+1)$, and $\ln x$ vs x .

Some of the more precise measurements made at this Laboratory¹³ are equivalent to a precision uncertainty of 6×10^{-7} cal/mole-deg which corresponds to an uncertainty in ρW of

$$\delta \rho = \frac{6 \times 10^{-7}}{k} = 1.8 \times 10^{17} .$$

The entropy is rarely determined with an accuracy of 1×10^{-3} cal/mole-deg, indicating that it is usually unnecessary to get the decimal point in W to better than about 10^{20} places from the correct location to compare the theory with the best experiments.²

Using Stirling's approximation, we have for the logarithm of the number of equally probable ways in which a given energy distribution may be achieved according to Boltzmann statistics:

$$\rho W_B = \rho N' + \sum_i (N_i \rho \epsilon_i - N_i \rho \gamma N_i + N_i) .$$

The variables N_i are not entirely independent, but are subject to the restrictions that total energy and total number of particles in the system be constant; i.e.,

$$\sum_i N_i = N = \text{constant} ,$$

and

$$\sum_i \epsilon_i N_i = E = \text{constant} .$$

These restrictions could be imposed upon the system before maximizing ρW_B by substituting for two of the N_i 's and reducing the number of variables by 2 using these last equations. However, Lagrange's method of undetermined multipliers can be applied so that all N_i 's are used. This preserves the symmetry of the equations and yields a relatively simple expression early in the problem.

Normally, when one seeks a stationary value of a function, F , of n variables, $x_1, x_2, \dots, x_n$, this value is determined by the condition that the variation

$$\delta F = \sum_{i=1}^n \frac{\partial F}{\partial x_i} \delta x_i = 0$$

for any conceivable small variation $\delta x_1, \dots, \delta x_n$ in F . That is,

$$\frac{\partial F}{\partial x_i} = 0$$

for all i .

If G is also a function of $x_1, \dots, x_n$, a substitution can be made for one of the x_i variables, say x_n , so that

$$\delta F' = \sum_{i=1}^{n-1} \frac{\partial F'}{\partial x_i} \delta x_i + \frac{\partial F'}{\partial G} \delta G ,$$

where F is now considered to be a function, F' , of $x_1, x_2, \dots, x_{n-1}, G$. If G is held constant, then $\delta G = 0$ and F' is maximized when $\partial F' / \partial x_i = 0$ for all i from 1 to $n-1$. At the point where these conditions are met, if G were not held constant, any variation in G would yield a variation in F' according to

$$\delta F' = \left. \frac{\partial F'}{\partial G} \right|_{x_1^0, \dots, x_{n-1}^0, G^0} \delta G .$$

But

$$\delta G = \sum_{i=1}^n \frac{\partial G}{\partial x_i} \delta x_i ,$$

so

$$\begin{aligned} \delta F &= \delta F' = \sum_{i=1}^n \left. \frac{\partial F'}{\partial G} \right|_{x_1^0, \dots, x_{n-1}^0, G^0} \frac{\partial G}{\partial x_i} \delta x_i \\ &= \sum_{i=1}^n \alpha \frac{\partial G}{\partial x_i} \delta x_i , \end{aligned}$$

with

$$\alpha \equiv \left. \frac{\partial F'}{\partial G} \right|_{x_1^0, \dots, x_{n-1}^0, G^0} .$$

Thus, at that point where F' is maximized with G held constant, the function F of $x_1, \dots, x_n$ satisfies

$$\delta F = \sum_{i=1}^n \left. \frac{\partial F'}{\partial G} \right|_{x_1^0, \dots, x_{n-1}^0, G^0} \frac{\partial G}{\partial x_i} \delta x_i ,$$

or

$$\begin{aligned}\frac{\partial F}{\partial x_i} &= \left. \frac{\partial F}{\partial G} \right|_{x_1^0, \dots, x_{n-1}^0, G^0} \frac{\partial G}{\partial x_i} \\ &= \alpha \frac{\partial G}{\partial x_i} \text{ with } G \text{ not held constant.}\end{aligned}$$

There are, therefore, two methods that may be used to find this point. In the first, G is held constant and F is maximized with respect to the remaining $n-1$ independent variables. In the second, G is not held constant, and the point is found that satisfies the conditions

$$\frac{\partial F}{\partial x_i} = \alpha \frac{\partial G}{\partial x_i} \text{ for all } i.$$

This adds the derivative α to the x_i^0 unknowns, giving $n+1$ unknowns and requiring $n+1$ equations for complete solution. α is known as the undetermined multiplier and is most often introduced as the requirement that

$$\delta F = \delta(F - \alpha G) = \sum_{i=1}^n \left(\frac{\partial F}{\partial x_i} - \alpha \frac{\partial G}{\partial x_i} \right) \delta x_i = 0.$$

The extension to two conditions follows similar lines and is not discussed in detail.

Using Lagrange's method, we therefore maximize the function $\ln W_B$ subject to constant N and E by maximizing $(\ln W_B - \alpha N - \beta E)$ where α and β are undetermined multipliers. Thus

$$\begin{aligned}\frac{\partial (\ln W_B - \alpha N - \beta E)}{\partial N_i} &= \frac{\partial}{\partial N_i} (N_i \ln g_i - N_i \ln N_i \\ &\quad + N_i - \alpha N_i - \beta \epsilon_i N_i) \\ &= \ln g_i - \ln N_i - 1 + 1 - \alpha - \beta \epsilon_i \\ &= \ln \frac{g_i}{N_i} - \alpha - \beta \epsilon_i = 0\end{aligned}$$

$$\ln \frac{g_i}{N_i} = \alpha + \beta \epsilon_i$$

$$N_i = g_i e^{-\alpha - \beta \epsilon_i},$$

which is a relatively simple expression, with the multipliers α and β , however, remaining to be determined. At very high energies, the exponential term in this equation conflicts with the actual energy distribution of a real finite system because no real distribution can ever allow any particle to occupy an energy level higher than the maximum energy available to the system. This restriction is necessitated by the First Law of Thermodynamics. The exponential distribution derived above does indicate a finite fractional occupancy for energy levels $\epsilon_i > E$. This is because using Stirling's approximation for $\ln N_i!$ involves using a continuous function that is not restricted to integer values. Consequently, maximizing $\ln W_B$ without restriction to integer values allows continuous values for N_i including values between zero and unity. The term $N_i \epsilon_i$ remains less than E even when $\epsilon_i > E$ if N_i is given a small enough fractional value. Fractional values can be obtained for real systems by time averaging. However, each individual distribution included in the time average will have zero occupancy of energy levels above E . Therefore, the time-averaged distributions obtained for real systems will be as nearly continuous as quantization restrictions allow, but will indicate zero probability at energies above E , in contrast to a finite value assigned by the Boltzmann distribution. This discrepancy lies above the energy range that is usually of interest, however. More will be said about this matter later.

The distribution law for the translational energy of an ideal gas containing N molecules, each of mass m can be obtained readily. The energy of any molecule in such a system depends only upon its speed.

$$\epsilon = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

The number of molecules, dN , with velocity components in the ranges $\dot{x}$ to $\dot{x} + d\dot{x}$, $\dot{y}$ to $\dot{y} + d\dot{y}$, and $\dot{z}$ to $\dot{z} + d\dot{z}$ is proportional to $e^{-\beta \epsilon}$ and to the magnitude of the velocity range because of equal likelihood assigned to equal ranges of momentum, giving

$$dN = K e^{-\frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)\beta} d\dot{x}d\dot{y}d\dot{z},$$

where K is a constant of proportionality. Carrying out the integration gives

$$\begin{aligned} N &= K \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)\beta} d\dot{x}d\dot{y}d\dot{z} \\ &= K \left(\frac{2\pi}{m\beta} \right)^{3/2}, \end{aligned}$$

or

$$K = N \left(\frac{m\beta}{2\pi} \right)^{3/2}.$$

So, in polar coordinates,

$$dN = N \left(\frac{m\beta}{2\pi} \right)^{3/2} e^{-\frac{m}{2}\beta c^2} \sin \theta dc d\theta d\phi.$$

The number of molecules, dN_c , in the range c to $c + dc$ is

$$\begin{aligned} dN_c &= \int dN \\ &= N \left(\frac{m\beta}{2\pi} \right)^{3/2} \int_0^{2\pi} \int_0^\pi e^{-\frac{m}{2}\beta c^2} \sin \theta d\theta d\phi \\ &= N \left(\frac{m\beta}{2\pi} \right)^{3/2} e^{-\frac{m}{2}\beta c^2} 4\pi c^2 dc. \end{aligned}$$

$$\text{Because } \epsilon = \frac{1}{2} m c^2,$$

$$dN_\epsilon = \frac{2N}{\sqrt{\pi}} \beta^{3/2} e^{-\beta\epsilon} \sqrt{\epsilon} d\epsilon.$$

For one mole of gas, N is Avogadro's number and

$$\begin{aligned} E &= \int_0^\infty \epsilon dN_\epsilon \\ &= \int_0^\infty \frac{2N}{\sqrt{\pi}} \beta^{3/2} e^{-\beta\epsilon} \epsilon^{3/2} d\epsilon \\ &= \frac{3}{2} \frac{N}{\beta}. \end{aligned}$$

But for one mole of an ideal gas, $E = \frac{3}{2}$ and $RT = \frac{3}{2} NkT$. Therefore, $\beta = 1/kT$. The energy distribution law is then

$$dN_\epsilon = \frac{2N}{\sqrt{\pi} (kT)^{3/2}} e^{-\frac{\epsilon}{kT}} \sqrt{\epsilon} d\epsilon.$$

For any given finite, isolated system, there is some maximum kinetic energy of translation that can be achieved by a single particle of a given type. If a gas having very little energy stored as potential energy is considered, then the maximum kinetic energy may be nearly ϵ_{kin} , the average total kinetic energy of the gas molecules.

For the hard-sphere model used to illustrate the distributions of systems containing various numbers of particles, one further restriction prevails. The particles are initially given velocities such that the net momentum for the hard-sphere molecules is zero. Because periodic boundaries are used, there is no interaction of the particles with the walls of a container. For this special case, the law of conservation of momentum requires that the net momentum for the molecules remain at zero. If the molecules had, instead, been allowed to reflect off the walls of a very massive container, the net momentum for the molecules themselves would not have been restricted to the initial value of zero.

For the hard-sphere system with no net momentum relative to the reference axes, the maximum energy obtainable by any particle is $E - \epsilon_{\text{av}}$, where ϵ_{av} is the average energy per molecule and E is the total kinetic energy. This may be shown by maximizing the energy of the particle of interest, say particle ℓ , while holding the total energy constant and requiring the total momentum to be zero. The momenta of the particles other than particle ℓ must be directed opposite to the momentum of particle ℓ so that a minimum of energy given to them will bring the net momentum of the system to zero. Thus, we require that

$$\frac{\partial}{\partial \epsilon_j} \left(\epsilon_\ell - a \sum_{k=1}^N \epsilon_k - b \sqrt{2m} \sum_{k=1}^N \sqrt{\epsilon_k} \right) = 0,$$

where ϵ_k represents the energy of particle k . Note that here the subscripts refer to the individual particles, not to the individual energy levels as in the derivation of the Boltzmann distribution. The summations run from 1 up to N where N is a constant representing the total finite number of molecules contained in the isolated system. Consequently, the system is also subject to the constraint of constant N by the nature of the summation and by the fact that N is not allowed to vary with ϵ_j .

For $j \neq l$, then

$$-a - \sqrt{\frac{m}{2\epsilon_j}} b = 0 ,$$

$$\sqrt{\epsilon_j} = -\sqrt{\frac{m}{2}} \frac{b}{a} ,$$

$$\epsilon_j = \frac{m}{2} \left(\frac{b}{a} \right)^2 .$$

For $j = l$

$$1 - a - \sqrt{\frac{m}{2\epsilon_l}} b = 0 ,$$

$$\sqrt{\epsilon_l} = \sqrt{\frac{m}{2}} \frac{b}{1-a} ,$$

$$\epsilon_l = \frac{m}{2} \left(\frac{b}{1-a} \right)^2 .$$

Because the total momentum for the system is zero,

$$\sum_{k=1}^N \sqrt{2m\epsilon_k} = 0 ,$$

or

$$\sqrt{2m\epsilon_l} + \sum_{k=1}^N{}' \sqrt{2m\epsilon_k} = 0 ,$$

where the prime indicates that the summation does not include $k = l$.

$$\frac{mb}{1-a} - (N-1) \frac{mb}{a} = 0 ,$$

where the factor $(N-1)$ appears because there are $(N-1)$ terms in the primed summation.

$$\frac{1}{1-a} = \frac{N-1}{a} ,$$

and

$$a = \frac{N-1}{N} .$$

Because the total energy of the system is E ,

$$\epsilon_l + \sum_{k=1}^N{}' \epsilon_k = E ,$$

or

$$\frac{m}{2} b^2 \left[\frac{1}{(1-a)^2} + \frac{N-1}{a^2} \right] = E ,$$

$$\frac{m}{2} b^2 \frac{N^3}{N-1} = E ,$$

$$\frac{mb^2}{2} = \frac{N-1}{N^2} \frac{E}{N} = \frac{N-1}{N^2} \epsilon_{av} .$$

Therefore,

$$\epsilon_l = \frac{mb^2}{2} \frac{1}{(1-a)^2}$$

$$= \frac{N-1}{N^2} \epsilon_{av} N^2$$

$$= (N-1) \epsilon_{av}$$

$$= E - \epsilon_{av} .$$

And, of course,

$$\epsilon_j = \frac{\epsilon_{av}}{N-1} \text{ for } j \neq l .$$

Had the momentum restriction not been placed upon the system, the maximum energy obtainable by one of the hard-sphere molecules would have been

the total kinetic energy, E , available to the system.

No energy would be available for the other particles, so that in this state $\epsilon_l = E$ and $\epsilon_j = 0$ for each of the other molecules.

For the case involving the momentum restriction, one can consider a series of isolated systems each having the same total energy, E . Those systems containing more molecules will have lower values for ϵ_{av} than will systems with fewer molecules. Consequently, the upper limit of energy available to any given molecule will be closer to E in isolated systems containing many molecules than in systems with fewer molecules.

IV. COMPUTER MODEL

To represent a macroscopic system as nearly as possible by the few molecules that can reasonably be handled by computer, we used periodic boundaries. The molecules were contained in a cubic box with penetrable walls, but constrained so that as any given molecule left the box through one wall it reentered with unchanged velocity through the opposite wall. This eliminates all interaction of the particles with the walls and leaves only molecule-molecule interactions occurring within the system.

To simplify the calculations and so improve the speed of the program, we chose a simple hard-sphere model for each molecule. The potential energy, V , for this model is

$$V = \infty, r \leq 2$$

$$V = 0, r > 2,$$

where r is the magnitude of the separation of centers of a pair of molecules and the effective radius of each molecule is taken as one arbitrary unit of distance. This model allows the events in a many-bodied system to consist of a series of two-body collisions because each molecule changes velocity only at the instant when it is separated from another particle by the collision interatomic distance of two units. For a small number of particles, there is insignificant likelihood that more than one such collision will occur within the smallest interval of time that the computer can handle.

Initially, all molecules were given equal kinetic energies with their velocities having random directions subject to the restriction that the total momentum of the system be zero. Once the initial configuration has been set up, the computer calculates the time at which the first collision occurs. For these calculations, we used an IBM 7094 digital computer that carries 9 significant digits that can multiply 10 to a power ranging from 38 to -38. To determine which two molecules will be involved in the most immediate collision to take place, we calculate the time required for each pair traveling along their projected paths to reach a separation of 2, if possible. The smallest such time with a positive sign is taken as the time for the next collision. The periodic boundary conditions must be considered when calculating all possible collision times. The next collision is then calculated.

After changing the velocities of the two molecules involved, we added the time that each molecule had spent in its previous energy state to the total time that all molecules spent in a similar energy state. The program then returned to examining the system to determine the next event. Details of the calculations are given in reference 14.

Upon stopping this part of the program, we determined an energy distribution for the system by calculating densities for small increments of energy. Each density was calculated by dividing the total time spent by particles in a given energy range by the time the system had been running and also by the width of the energy increment to which the density applied. The widths of the increments were varied so that the total number of times particles had fallen into any given energy range was the same as that for any other range up to a cutoff value of 20 energy units.

The energy distribution actually obtained from the molecules was compared with the previously derived Boltzmann energy distribution.

$$\frac{dN}{d\epsilon} = \frac{2N}{\sqrt{\pi}(kT)^{3/2}} e^{-\frac{\epsilon}{kT}} \sqrt{\epsilon}.$$

Because each particle was initially given one unit of energy, and the average energy per molecule is related to the temperature by

$$\bar{\epsilon} = 3/2 kT \quad ,$$

kT must have the value 2/3. So with N representing the number of molecules in the system,

$$\frac{dN}{d\epsilon} = \frac{2N}{\sqrt{\pi}} \left(\frac{3}{2}\right)^{3/2} e^{-\frac{3}{2}\epsilon} \sqrt{\epsilon}$$

is the function plotted against ϵ as a solid line in the figures that follow.

V. RESULTS

To prevent the initial condition that every molecule have one unit of energy from distorting the energy distribution around unity, we allowed each system examined to run for some time to achieve a more likely energy configuration before starting the count for the distribution calculations.

We examined six different systems containing 4, 4, 8, 16, 32, and 108 molecules, respectively. In every case but the third and fourth, the initial configuration was face-centered cubic. For instance, in the first system with a container edge 12 units long, the 4 molecules were initially placed as follows.

0, 0, 0
6, 6, 0
6, 0, 6
0, 6, 6

However, because of the periodic boundary conditions, the computer had to consider the further positions

(-6, 6, 0)
(6, -6, 0)
(-6, -6, 0)
(-6, 0, 6)
(6, 0, -6)
(-6, 0, -6)

(0, -6, 6)
(0, 6, -6)
(0, -6, -6) .

With a 12-unit edge, the first system had an initial nearest-neighbor interatomic distance of 8.48 units which allowed the molecules to pass by each other easily so that the initial configuration was soon lost. Also, because of this, the molecules were able to pass through the walls. The 2nd, 5th, and 6th systems with container edges of 3, 6, and 9 units, respectively, had nearest-neighbor intermolecular distances of 2.12 which is only 0.12 unit greater than the collision diameter of 2. In these three systems, there was insufficient room for the molecules to pass each other, so they prevented each other from passing through the container walls. The remaining systems were not so restricted. The numerical results for the various systems are given in Table II.

Figure 2 compares the results for four molecules in a container large enough to allow random wandering with the results from a container small enough to restrict the molecules to roughly their original face-centered positions. The results are essentially the same, which agrees with the theory.

Because the first two systems each contain four molecules, the total energy in either system is four energy units. From previous discussion, the maximum energy that can be obtained by any molecule while the net system momentum is kept at zero is $E - \epsilon_{av}$. The maximum energy range for the measured distribution must be 4 - 1, or 3, units. Although time-averaging allows a continuous distribution to be obtained even with only four molecules, the fact that the upper energy limit is very close to the average energy (three times greater) causes moderate distortion of the distribution from that of Boltzmann. The total area under the measured distribution must be the same as that under the Boltzmann distribution because this area represents the number of particles in the system. For the measured distribution, this area must be contained within the range from 0 to 3 rather than from 0 to ∞ ; hence the distortion. Nevertheless, even for only four molecules, a Boltzmann-like distribution begins to appear.

TABLE 11
NUMERICAL RESULTS FOR THE VARIOUS SYSTEMS

<u>System 1</u>	<u>System 2</u>	<u>System 3</u>	<u>System 4</u>	<u>System 5</u>	<u>System 6</u>
Number of Molecules					
4	4	8	16	32	108
Length of Container Edge					
12	3	6	6	6	9
Number of Times Container Walls Were Penetrated					
231,677	0	29,212	9,872	0	0
Number of Collisions					
33,997	158,465	57,789	68,736	23,112	53,017
Sampling Time Interval					
4.89×10^5	5.40×10^3	1.56×10^4	2.67×10^3	8.74×10^1	5.89×10^1
Number of Events Recorded in Each Energy Range					
5,000	20,000	5,000	4,000	4,000	4,000
Density vs Energy					
1.54 0.13	1.47 0.12	3.15 0.07	6.35 0.05	16.61 0.11	46.12 0.06
2.30 0.40	2.25 0.30	4.88 0.20	9.93 0.14	22.84 0.30	70.63 0.17
2.48 0.49	2.45 0.44	5.36 0.30	10.45 0.21	22.54 0.45	76.61 0.24
2.48 0.63	2.46 0.56	5.35 0.38	10.94 0.27	20.13 0.60	79.49 0.32
2.48 0.76	2.50 0.68	5.28 0.46	11.66 0.32	18.46 0.76	77.74 0.38
2.29 0.90	2.40 0.80	5.48 0.54	11.33 0.38	16.56 0.92	74.89 0.46
2.20 1.04	2.31 0.92	5.12 0.62	11.24 0.43	13.43 1.11	73.93 0.52
2.01 1.17	2.21 1.04	4.91 0.70	10.79 0.48	10.70 1.33	71.99 0.59
1.88 1.32	2.03 1.16	4.60 0.78	10.56 0.54	8.02 1.60	66.54 0.66
1.61 1.48	1.90 1.28	4.45 0.86	10.72 0.60	4.90 1.98	63.84 0.73
1.28 1.66	1.68 1.42	4.16 0.96	10.12 0.65	1.96 2.70	59.13 0.81
0.97 1.88	1.46 1.56	3.83 1.04	9.94 0.70		55.60 0.89
0.49 2.24	1.18 1.73	3.48 1.14	9.26 0.76		53.04 0.97
	0.88 1.94	3.22 1.24	8.76 0.82		47.62 1.06
	0.49 2.24	2.78 1.36	8.25 0.88		43.73 1.14
		2.52 1.48	8.09 0.94		38.80 1.24
		2.14 1.62	7.67 1.00		35.44 1.34
		1.78 1.78	7.14 1.06		31.08 1.46
		1.40 1.96	6.72 1.14		25.07 1.59
		0.99 2.20	6.14 1.20		22.39 1.74
		0.58 2.54	5.45 1.28		17.76 1.90
			5.23 1.36		14.04 2.12
			4.72 1.44		9.66 2.39
			4.34 1.54		5.66 2.80
			3.70 1.64		1.58 3.86
			3.22 1.76		
			2.82 1.88		
			2.43 2.03		
			1.94 2.20		
			1.39 2.43		
			0.88 2.74		
			0.42 3.28

Fig. 2. Energy distribution of four molecules compared with the Boltzmann distribution.

As the number of molecules contained in the system is increased, the measured distribution appears to approach that predicted by theory, as is shown in Figs. 3 through 6. The final system, which contains 108 molecules (an extremely small number compared to most real systems), has a measured distribution that seems to agree very well with that of Boltzmann.

VI. CONCLUSIONS

The hard-sphere model has allowed simple specification of exact equations governing the motions of a moderate number of particles. These equations are reversible with respect to time, and the be-

Fig. 3. Energy distribution of eight molecules compared with the Boltzmann distribution.

Fig. 4. Energy distribution of 16 molecules compared with the Boltzmann distribution.

havior of the system as a whole at equilibrium appears to be, also, for this assumption is implicit in the development of Boltzmann statistics. The assumption of equal a priori probability for equal extensions in phase space appears to be valid for such systems.

REFERENCES

1. Frederick T. Wall, Chemical Thermodynamics, W. H. Freeman and Co., San Francisco, 1958, p. 228.
2. G. N. Lewis, M. Randall, K. S. Pitzer, and L. Brewer, Thermodynamics, McGraw-Hill Book Company, New York, 1961, p. 635.

Fig. 5. Energy distribution of 32 molecules compared with the Boltzmann distribution

Fig. 6. Energy distribution of 108 molecules compared with the Boltzmann distribution.

3. K. S. Pitzer, Quantum Chemistry, Prentice-Hall, Inc., Englewood Cliffs, 1953, p. 100.
4. J. E. Mayer and M. G. Mayer, Statistical Mechanics, John Wiley & Sons, Inc., New York, 1940, p. 109.
5. R. C. Tolman, The Principles of Statistical Mechanics, Oxford University Press, London, 1938, pp. 65, 71.
6. R. W. Hakala, J. Chem. Educ. 38, 33 (1961).
7. R. W. Hakala, J. Chem. Educ. 39, 526 (1962).
8. E. G. D. Cohen and T. H. Berlin, Physica 26, 717 (1960).
9. M. S. Green, J. Chem. Phys. 25, 836 (1956).
10. P. Mazur and J. Biel, Physica 32, 1633 (1966).
11. R. F. Snider, J. Chem. Phys. 41, 591 (1964).
12. E. G. D. Cohen, Physica 27, 163 (1961).
13. D. R. Conant, J. Chem. Eng. Data 16, 430 (1971).
14. D. R. Conant, Los Alamos Scientific Laboratory Report LA-4264 (1970).