

A Simple Derivation of the Maxwell-Boltzmann Law

E. U. CONDON

Westinghouse Research Laboratories, East Pittsburgh, Pennsylvania

(Received August 11, 1938)

The Maxwell-Boltzmann law is derived in a direct and simple way from the usual postulates that the mechanical system has a discrete system of allowed states and that each of these states has equal *a priori* weight in calculation of statistical averages.

THE Maxwell-Boltzmann distribution law plays such a fundamental role in the entire theory of thermal phenomena that it is important to have as simple an approach to it as possible. In this paper, a method of derivation is presented which is believed to be essentially new¹ and which shows in an extremely simple manner the use of the assumptions involved.

There are only two basic assumptions needed: (1) The dynamical system in question is governed by quantum mechanics and, being a closed system, has, therefore, a discrete spectrum of allowed energy levels, (2) in calculating statistical averages, each state corresponding to one linearly independent wave function orthogonal to all the others, that is, consistent with known features of the problem, is to be given the same weight.

Let us first consider the case of a large number N of noninteracting molecules in a cubical box of edge L and volume $V=L^3$. The wave function for an allowed state of a single molecule is

$$u_{lmn}(x, y, z) = \left(\frac{2}{V}\right)^{\frac{1}{2}} \sin \frac{l\pi x}{L} \sin \frac{m\pi y}{L} \sin \frac{n\pi z}{L} \quad (1)$$

and the corresponding energy is

$$W(l, m, n) = (h^2/8\mu L^2)(l^2 + m^2 + n^2). \quad (2)$$

The allowed states for N noninteracting mole-

cules in a box will have as wave functions a continued product of such one-particle wave functions (1) and the total energy of the N particles will be a sum of such one-particle energy expressions (2). Neglecting the symmetry restrictions on the wave function which would give rise to Einstein-Bose and Fermi-Dirac statistics, there will be one state of the system for each possible complete set of quantum numbers consisting of N individual sets $(l_\alpha, m_\alpha, n_\alpha)$, where $\alpha=1, 2, \dots, N$ and the $3N$ quantum numbers $l_\alpha, m_\alpha, n_\alpha$ range independently over all positive integral values.

It is important to know the total number of states of the system whose total energy is equal to or less than W for values of W large compared to the interval between energy levels. This is found by the method that has often been used in statistical mechanics. We introduce a $3N$ dimensional space. In this space the state of the system whose quantum numbers are $l_1 m_1 n_1, l_2 m_2 n_2, \dots, l_N m_N n_N$ is associated with the point having this set of positive integers for Cartesian coordinates. There is thus one state per unit volume in this space. The number of states for which the energy is less than W is thus equal to 2^{-3N} times the volume of a $3N$ dimensional sphere of radius $(8\mu L^2 W/h^2)^{\frac{1}{2}}$.

Denoting the number of states for the N particle system whose energy is less than W by $C_N(W)$, it is evident from dimensional considerations alone that $C_N(W)$ is proportional to $W^{3N/2}$ which fact is all that is needed in the derivation of the distribution law to be given below. From geometry it is known that the volume of an n dimensional sphere of radius r is

$$(\pi^{n/2} r^n) / \Gamma\left(\frac{n}{2} + 1\right) \quad (3)$$

¹ *Note added in proof:*—Since this was written, I have noticed that a similar derivation is given by Kennard in his new book *Kinetic Theory of Gases*, p. 390. I have also had an interesting discussion with Professor G. E. Uhlenbeck, who remarks that the approach followed here is essentially a modern version of part of Maxwell's work as presented, for example, in Jeans' *Dynamical Theory of Gases*, Chapter V, p. 119 *et seq.* He also remarked on its connection with the discussion of statistical distribution of energy among a small number of particles as worked out for a problem in nuclear physics in a paper by Uhlenbeck and Goudsmit, pp. 201–211 in the *Zeeman Verhandelingen* (Martinus Nijhoff, The Hague, 1935).

and, therefore, the exact formula for $C_N(W)$ is

$$C_N(W) = \frac{1}{\Gamma\left(\frac{3N}{2} + 1\right)} \left[\frac{(2\pi\mu W)^{\frac{1}{2}} L}{h} \right]^{3N} \quad (4)$$

Now let us consider a system consisting of $(N+1)$ noninteracting molecules in the box and ask for the probability that the extra molecule have an energy between w and $w+dw$ when all we know of the system is that the total energy is W . The probability, according to the statistical postulate, will be proportional to the number of states of the composite system for which w lies between w and $w+dw$ and hence the energy w_N of the other N molecules lies between $W-w$ and $W-w-dw$. Therefore, writing $P(w)dw$ for the probability that the energy of the extra molecule lie in this range, we have

$$P(w)dw \sim C_N'(W-w)C_1'(w)dw, \quad (5)$$

where $C_N'(w)$ is the derivative of $C_N(w)$ with respect to w . Using only that part of (4) which makes $C_N(W)$ proportional to $W^{3N/2}$, we have

$$P(w)dw \sim (1-w/W)^{3N/2-1} w^{\frac{1}{2}} dw. \quad (6)$$

Let us define an amount of energy called kT by the equation

$$W = 3NkT/2. \quad (7)$$

Then, since N is very large, we can recognize that the first factor is equal to the exponential factor of the Boltzmann distribution, so

$$P(w)dw \sim e^{-w/kT} w^{\frac{1}{2}} dw, \quad (8)$$

which is the usual Maxwellian distribution of velocities.

This mode of derivation brings out clearly that the $w^{\frac{1}{2}}$ factor arises from the fact that there are more states of the single molecule available in unit energy range at higher energies than at lower energies, whereas the exponential factor arises from the fact that if the single molecule gets more energy, there is necessarily less left for the other N molecules, and thus they are required to be in a range of energy where fewer states per unit energy range are available for them.

It is now a simple calculation to normalize (8) and find that

$$P(w)dw = (2/\pi^{\frac{1}{2}}) e^{-w/kT} (w/kT)^{\frac{1}{2}} d(w/kT) \quad (9)$$

and to calculate that the mean energy of a single molecule is

$$w_{av} = 3kT/2. \quad (10)$$

Physically, the N molecules in the composite system can be regarded as the perfect gas thermometer with which the single molecule is in thermal equilibrium.

Next, we can consider a slight generalization of the foregoing discussion, which leads to the Boltzmann distribution for systems having such a widely spaced set of allowed levels that they cannot be handled by means of a continuous $C(w)$ function.

We consider a composite system as before, which consists of N molecules in a box and in addition, the arbitrary quantized system whose allowed energy levels will be written $w_1, w_2, \dots, w_\alpha, \dots$ with the corresponding statistical weights (order of degeneracy) $g_1, g_2, \dots, g_\alpha$. We suppose that the total energy of the composite system is known to lie between W and $W+\delta W$ and ask for the probability that the quantized part be found in the α th energy level, assuming that all states of the composite system with total energy between W and $W+\delta W$ are equally probable.

The probability of finding the quantized part in the α th energy level will, therefore, be proportional to

$$P(W_\alpha) \sim C_N'(W-w_\alpha)g_\alpha\delta W,$$

which is also

$$P(w_\alpha) \sim (1-w_\alpha/W)^{3N/2-1} g_\alpha \delta W.$$

As before, we may introduce the energy kT defined by (7) and recognize that the first factor is essentially equal to $e^{-w_\alpha/kT}$ if N is a very large number. Therefore, we have

$$P(w_\alpha) \sim g_\alpha e^{-w_\alpha/kT}, \quad (11)$$

which is the familiar Boltzmann distribution law for systems having quantized energy levels. As before, we recognize the fact that the probability is proportional to g_α , the number of states of the quantized system of energy w_α , and the exponential factor representing the dependence on w_α of the density of states of the N molecules in a box which constitute the perfect gas thermometer-thermostat with which the quantized system is in equilibrium.

With the derivation of the Boltzmann law accomplished, one can proceed to develop the theory of the thermal properties of matter in the usual way by introducing the partition function

$$z(T) = \sum_{\alpha} g_{\alpha} e^{-w_{\alpha}/kT}, \quad (12)$$

from which the mean energy at temperature T is calculated by the formula

$$\bar{w}(T) = -k \frac{d \log z}{d(1/T)}. \quad (13)$$

The uses of the partition function are so well known that it will not be necessary here to repeat that part of the development.

In deriving the distribution law for the distribution of translational energy of a single molecule in equilibrium with N molecules, the assumption was made that the energy of the $(N+1)$ molecules is precisely known to be W . Looking back over the argument leading to (8), we see that the argument would have been essentially unaltered had we assumed that the composite system's energy had a value between W and $W + \delta W$.

Actually, when we have a gas at temperature T it will have a distribution-in-energy of its total energy, which we may derive by assuming the N molecules in a box to be part of a much larger composite system which includes N' more molecules where N' is very large compared to N . In that case we find that the distribution in energy of the N molecules is given by

$$P(w)dw \sim e^{-w/kT} w^{3N/2-1} dw, \quad (14)$$

where kT is defined by $W = 3N'kT/2$ which is supposed to be negligibly different from $3(N' + N)kT/2$ since $N' \gg N$. From (14) we may calculate the mean energy of the system of N molecules in thermal equilibrium with the larger system. It comes out

$$\bar{w} = 3NkT/2, \quad (15)$$

which justifies the identification of T with the usual absolute temperature on the perfect gas scale. Similarly one may calculate the mean-square deviation from the mean of the energy of the N molecules

$$\Delta^2 = (\bar{w} - w_{Av})^2_{Av} = 3N(kT)^2/2. \quad (16)$$

Hence the fractional fluctuation in energy of the

N molecules is

$$\Delta/w_{Av} = (3N/2)^{-1/2}, \quad (17)$$

which is extremely small when N is large. This justifies the neglect of the distribution in energy of the N molecules in the argument leading to (8) by which the distribution in energy of one molecule was found.

The distribution law for the case of Fermi-Dirac statistics² is easily obtained as follows. We consider the system of N equivalent particles to which the Fermi statistics is to be applied, as a single quantized system in equilibrium with a larger perfect gas-thermometer thermostat. Then the distribution in energy of its quantized states is given by application of (11). In other words the relative probability of each of the independent states of the N equivalent particles governed by the Pauli exclusion principle is given by the Boltzmann factor, $e^{-w/kT}$.

The exclusion principle tells us that no two particles can be in the same quantum state and, therefore, if the allowed energy levels of a single particle are given by w_{α} , each allowed level of the system of N particles will be characterized by a set of quantum numbers N_{α} , one for each single particle state α , where N_{α} can have only the values 0 and 1 and $\sum_{\alpha} N_{\alpha} = N$. The total energy is

$$W = \sum_{\alpha} w_{\alpha} N_{\alpha}.$$

To find the distribution-in-energy of the single particle in the case of Fermi-Dirac statistics, we have to calculate the mean value of N_{β} , the probability of occupation of the β th state by a particle. By the Boltzmann principle this is

$$(N_{\beta})_{Av} = \frac{\sum_0 e^{-W/kT} + \sum_1 e^{-W'/kT}}{\sum_{0+1} e^{-W/kT}},$$

where \sum_0 means the sum over all sets of N 's consistent with $\sum N_{\alpha} = N$ and having $N_{\beta} = 0$ while \sum_1 means summation over all sets of N 's having $\sum N_{\alpha} = N$ and having $N_{\beta} = 1$. If we write

$$W' = W - N_{\beta} w_{\beta},$$

this can be written

$$(N_{\beta})_{Av} = \frac{e^{-w_{\beta}/kT} \sum_1 e^{-W'/kT}}{\sum_0 e^{-W'/kT} + e^{-w_{\beta}/kT} \sum_1 e^{-W'/kT}}$$

² A valuable account of Fermi-Dirac gas theory, together with its most important field of application, the Sommerfeld electron theory of metals, is given in the article by Sommerfeld and Bethe, *Handbuch der Physik*, Vol. 24/2 (Julius Springer, Berlin, 1933), p. 333.

and, therefore,

$$(N_\beta)_{Av} = \frac{1}{A(T)e^{w_\beta/kT} + 1}, \quad (18)$$

$$\text{in which } A(T) = \frac{\sum_0 e^{-w'/kT}}{\sum_1 e^{-w'/kT}}. \quad (19)$$

Here the parameter $A(T)$ is the ratio of the partition function for an N particle system from which the state w_β is excluded from the set of allowed single particle states to that for an $(N-1)$ particle system from which this same state is excluded.

Evidently, if $A(T) \gg 1$, then the $+1$ in the denominator of (18) will be negligible and $(N_\beta)_{Av}$ will be approximately equal to $e^{-w_\beta/kT}$ times a factor that is independent of w_β . Therefore, in this limit the distribution law is not appreciably affected by the operation of the Pauli exclusion principle. On the other hand, if $A(T) \ll 1$ then for values of w_β small enough that $A(T)e^{w_\beta/kT} \ll 1$, we shall have $(N_\beta)_{Av} = 1$, that is, the low energy states are almost certainly occupied by one particle in each state which means a large departure from the classical distribution law.

The probability that a single particle have energy between w and $w+dw$ is, therefore, pro-

portional to the product of the number of states in this energy range; $c'(w)dw$ multiplied by the chance that a single particle state of energy w be occupied which is $[Ae^{w/kT} + 1]^{-1}$ so the distribution function is

$$P(w)dw \sim c'(w)[Ae^{w/kT} + 1]^{-1}dw, \quad (20)$$

which is the familiar form from which the usual deductions of properties of the Fermi-Dirac gas may be made.

In conclusion, it is fitting to remark on the occasion of the seventieth birthday of one of the greatest teachers and productive workers that modern theoretical physics has known—Arnold Sommerfeld. Everyone of my generation grew up on atomic physics by way of his great *Atombau und Spektrallinien*, a large group have profited by the stimulation of his lectures on his American visits, and a fortunate few of us have derived boundless stimulation from the opportunity of working in his Institut für theoretische Physik in the former brighter days. All physicists join in wishing him a happy birthday and continued vigor with which to participate in the further developments of fundamental ideas to which he has contributed so extensively.

Transitions Between Levels Spaced Almost Continuously

ALFRED LANDÉ

Mendenhall Laboratory, Ohio State University, Columbus, Ohio

(Received June 16, 1938)

The transition probabilities between closely spaced energy levels, when calculated according to the usual perturbation methods, lead to apparent infinities if the perturbation is finite. A more appropriate method of approximation which disposes of the infinities is worked out. The resulting finite transition probabilities are identical with the finite parts of the usual expressions in first and second order, but additional finite terms appear in orders higher than the second.

THE probabilities of transition between the levels of a continuous or almost continuous energy spectrum can always be treated in an unambiguous way as far as the first-order (direct) transitions are concerned. But the second-order transitions (through *one* intermediate state) give rise to infinite transition probabilities in the case of a continuous spectrum. Usually one splits up the resulting terms, in a more or less arbitrary

way, into a finite part that is supposed to represent the physical facts, and an infinite part that is neglected without further justification. This procedure appears all the more ambiguous when one learns that different approximation methods lead to different convergent parts of diverging series. The omission of infinite terms seems to be based on the hope that the various infinite members would cancel one another if