*Article*

# Translation of Ludwig Boltzmann's Paper "On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium" Sitzungberichte der Kaiserlichen Akademie der Wissenschaften. Mathematisch-Naturwissen Classe. Abt. II, LXXVI 1877, pp 373-435 (Wien. Ber. 1877, 76:373-435). Reprinted in Wiss. Abhandlungen, Vol. II, reprint 42, p. 164-223, Barth, Leipzig, 1909

**Kim Sharp \* and Franz Matschinsky**

E. R. Johnson Foundation, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; E-Mail: matsch@mail.med.upenn.edu

\* Author to whom correspondence should be addressed; E-Mail: sharpk@mail.med.upenn.edu; Tel.: +1-215-573-3506; Fax: +1-215-898-4217.

**Abstract:** Translation of the seminal 1877 paper by Ludwig Boltzmann which for the first time established the probabilistic basis of entropy. Includes a scientific commentary.

## 1. Introduction

Barely any of Boltzmann's original scientific work is available in translation. This is remarkable given his central role in the development of both equilibrium and non-equilibrium statistical mechanics, his statistical mechanical explanation of entropy, and our understanding of the Second Law of thermodynamics. What Boltzmann actually wrote on these subjects is rarely quoted directly, his methods

are not fully appreciated, and key concepts have been misinterpreted. Yet his work remains relevant today. Witness, for example, the recent debate about Gibbs entropy *vs.* Boltzmann entropy [1–4]. The paper translated here (LB1877) occupies the high middle ground of his work, and was a key step in developing the fully probabilistic basis of entropy and the Second Law.



**Figure 1.** The Entropic Boltzmann, by P. L. Dutton, FRS.

LB1877 exemplifies several of Boltzmann's most important contributions to modern physics. These include

(1) The eponymous Boltzmann distribution, relating the probability of a state to the exponential of its energy scaled by the mean kinetic energy (temperature). This distribution is so pervasive in statistical mechanics one almost forgets that it had to be "discovered" in the first place. Deriving the correct distribution is absolutely necessary for Boltzmann's statistical mechanical definition of entropy, and it is derived twice, in Sections I and II, for discretized and continuous energy distributions, respectively. Boltzmann considers this point important enough to devote a whole section (IV) to apparently plausible ways to determine the probabilities of different state distributions which are physically incorrect: Either because they don't produce the right dispersion in kinetic energies between molecules, or because the distributions don't depend correctly on the average kinetic energy (temperature), and so are not scale invariant. Previous work by Maxwell and Boltzmann was based on mechanical laws of motion and particle interaction in gases. In contrast, Boltzmann's derivations here are much more general. They require only that particles can exchange kinetic energy, but they don't specify the mechanism. As Boltzmann predicted in the final sentences of this paper his approach was applicable not just to gases, but to liquids and solids of any composition. Indeed, the Boltzmann distribution has also passed almost unchanged into the quantum world. Even without its connection to entropy, the Boltzmann distribution is of remarkably wide ranging importance: explaining, amongst other phenomena, the density distribution of the atmosphere, cooling by evaporation, and Arrhenius type chemical kinetics.

(2) Much of the theoretical apparatus of statistical mechanics is developed with great clarity in Sections I and II of LB1877. Here Boltzmann uses a hierarchy of three levels of description. First there is the macro-state, the level of thermodynamic observables such as temperature and pressure. Second there is the level at which the energy or velocity components of each molecule are specified. He calls this a Komplexion, which we translate literally as complexion. Finally there is the intermediate level at which the number of molecules with each energy/velocity/position is specified, regardless of which molecules those are (Boltzmann's $w_0, w_1$ *etc.*). Boltzmann calls a particular set of $w$-values a Zustandeverteilung, translated as a "state distribution' or "distribution of states'. His terminology and usage of the three levels is incisive, in some ways superior to the two modern terms macro-state and micro-state: The latter is sometimes applied to both complexions and state distributions, eliding a distinction Boltzmann was careful to keep. Boltzmann then deploys the machinery of permutations, state counting *etc.* to determine the all important state distribution number: the number of complexions contained within any given state distribution, which he denotes by $\mathcal{P}$. Boltzmann then shows how to find the state distribution $(w_0^{max}, w_1^{max}, \ldots)$ with the largest number of complexions $\mathcal{P}^{max}$ subject to constraints on the temperature and number of molecules. Boltzmann's postulate is that $(w_0^{max}, w_1^{max}, \ldots)$ is the most probable state distribution, and that this corresponds to thermal equilibrium. Although Boltzmann introduces discrete energy levels in Section I as a device of mathematical convenience and considers them unphysical, he hedges his bets on this issue in Section II by showing that there is no material difference using continuous energies. This indifference to details like discretization demonstrates again the great generality of the statistical mechanical world view Boltzmann was forging.

(3) The statistical mechanical formulation of entropy, Boltzmann's third great contribution, comes in Section V of LB1877. Boltzmann shows that the statistical mechanical quantity he denotes by $\Omega$ (multiplied by 2/3) is equal to the thermodynamic quantity entropy (S) as defined by Clausius, to within an additive constant. Boltzmann called $\Omega$ the Permutabilitätmass, translated here literally as "permutability measure". Given the permutability measure's central role, it is worth reviewing its definition. A state distribution is specified by the number of molecules having velocity components within some small interval $u$ and $u + du, v$ and $v + dv, w$ and $w + dw$ for every $u, v,$ and $w$ value in the range $(-\infty, +\infty)$, and having position coordinates in each small volume $(x, x + dx), (y, y + dy), (z, z + dz)$, ranging over the total volume $V$. For each state distribution there is a number of possible complexions. One particular state distribution has the most complexions and so therefore is the most probable. $\Omega$ is given by the logarithm of the number of complexions for that state with the most complexions.

(4) In Section V of LB1877 Boltzmann also clearly demonstrates that there are *two* distinct contributions to entropy, arising from the distribution of *heat* (kinetic energy) and the distribution in *space* of atoms or molecules: compare his Equations (34) and (61) for $\Omega$. In the initial effort to understand the nature of entropy, Carnot, Clausius, Maxwell, Kelvin and others focused almost entirely on the contribution from heat. Boltzmann welded the spatial and thermal parts together into a coherent statistical mechanical formulation. It is fitting that Boltzmann was the one to discover the third fundamental contribution to entropy, namely radiation, by deriving the Stefan-Boltzmann Law [5].

Careful reading of LB1877 is enlightening with regard to a number of apparent paradoxes subsequently encountered in the development of statistical mechanics. First, regarding extensivity, $\Omega$

is not the logarithm of a probability. That would be obtained by dividing the number of complexions $\mathcal{P}$ for a given state distribution by the total number of complexions $J$ (see text immediately preceding Equation (1)); Boltzmann gives this a different symbol, W, from the first letter of the German word for probability, but he does not use it. Confusingly, Planck later chose to write Boltzmann's equation for entropy as S=klnW+constant [6]. Crucially, $\Omega$ is an extensive quantity, which leads to Boltzmann's extensive Equation (62) for the entropy of an ideal gas. Thus Boltzmann never encounters the apparent Gibbs paradox for the entropy of mixing of identical gases. Furthermore, with Boltzmann's Permutabilitätmass method for counting states, there is no need for *a posteriori* division by N! to "correct" the derivation using the "somewhat mystical arguments of Gibbs and Planck" [7] nor a need to appeal to quantum indistinguishability, which has been implausibly described as the appearance of quantum effects at the macroscopic classical level [8]. Subsequently, at least four distinguished practitioners of statistical mechanics have pointed out that correct counting of states *a la* Boltzmann obviates the need for the spurious indistinguishability/N! term: Ehrenfest [9], van Kampen [7], Jaynes [10], Swendsen [11] (and possibly Pauli [12]). This has had little impact on textbooks of statistical mechanics. An exception is the treatise by Gallavotti [13].

Second, regarding non-equilibrium states, $\Omega$ is not the logarithm of a volume in momentum-coordinate phase space d**V** = d**q**d**p** occupied by the system. Boltzmann notes, using Liouville's theorem, that d**V** remains constant in time (See Equation (41)), and so it cannot describe the entropy increase upon approach to equilibrium that Boltzmann was so concerned with. He thus avoids at the outset the considerable difficulty Gibbs had accounting for changes in entropy with time (see p144 in [14]) Boltzmann gave us, for the first time, a definition of entropy applicable to every state (distribution), at equilibrium or not: "Then the entropy of the initial and final states is not defined, (Tr: by Clausius' equation) but one can still calculate the quantity which we have called the permutability measure", LB1877, Section V. See also Equations (3), (31) and the definition of a state distribution as any set of $w's$ satisfying the constraints of number and mean kinetic energy. By extension, every complexion can then be assigned an entropy, using the permutability measure of the state distribution to which that complexion belongs [15], opening the door to the statistical mechanics of non-equilibrium states and irreversible processes.

Finally, for a more extensive appraisal of Boltzmann's entire oeuvre we recommend reading the superb biography of Boltzmann: "The Man Who Trusted Atoms" [16].

## 2. On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations regarding the Conditions for Thermal Equilibrium by Ludwig Boltzmann

The relationship between the second fundamental theorem and calculations of probability became clear for the first time when I demonstrated that the theorem's analytical proof is only possible on the basis of probability calculations. (I refer to my publication "Analytical proof of the second fundamental theorem of the mechanical theory of heat derived from the laws of equilibrium for kinetic energy" Wien.

Ber. 63, p8 reprinted as Wiss. Abhand. Vol I, reprint 20, pp295 [1] and also my "Remarks about several problems in the mechanical theory of heat"; 3rd paragraph, Wiss. Abhand. Vol II, reprint 39. This relationship is also confirmed by demonstrating that an exact proof of the fundamental theorems of the equilibrium of heat is most easily obtained if one demonstrates that a certain quantity—which I wish to define again as E [2]—has to decrease as a result of the exchange of the kinetic energy among the gaseous molecules and therefore reaching its minimum value for the state of the equilibrium of heat. (Compare my "Additional studies about the equilibrium of heat among gaseous molecules" Wiss. Abhand. Vol I, reprint 22, p316). The relationship between the second fundamental theorem and the laws of the equilibrium of heat is made even more compelling in light of the developments in the second paragraph of my "Remarks about several problems of the mechanical theory of heat". There I mentioned for the first time the possibility of a very unique way of calculating the equilibrium of heat using the following formulation: "It is clear that every single uniform state distribution which establishes itself after a certain time given a defined initial state is equally as probable as every single nonuniform state distribution, comparable to the situation in the game of Lotto where every single quintet is as improbable as the quintet 12345. The higher probability that the state distribution becomes uniform with time arises only because there are far more uniform than nonuniform state distributions"; furthermore: "It is even possible to calculate the probabilities from the relationships of the number of different state distributions. This approach would perhaps lead to an interesting method for the calculation of the equilibrium of heat." It is thereby indicated that it is possible to calculate the state of the equilibrium of heat by finding the probability of the different possible states of the system. The initial state in most cases is bound to be highly improbable and from it the system will always rapidly approach a more probable state until it finally reaches the most probable state, *i.e.*, that of the heat equilibrium. If we apply this to the second basic theorem we will be able to identify that quantity which is usually called entropy with the probability of the particular state. Let's assume a system of bodies which are in a state of isolation with no interaction with other bodies, e.g., one body with higher and one with lower temperature and one so called intermediate body which accomplishes the heat transfer between the two bodies; or choosing another example by assuming a vessel with absolutely even and rigid walls one half of which is filled with air of low temperature and pressure whereas the other half is filled with air of high temperature and pressure. The hypothetical system of particles is assumed to have a certain state at time zero. Through the interaction between the particles the state is changed. According to the second fundamental theorem, this change has to take place in such a way that the total entropy of the particles increases. This means according to our present interpretation that nothing changes except that the probability of the overall state for all particles will get larger and larger. The system of particles always changes from an improbable to a probable state. It will become clear later what this means. After the publication of my last treatise regarding this topic the same idea was taken up and developed further by Mr. Oskar Emil Meyer totally independent of me {Die kinetische Theorie der Gase; Breslau 1877, Seite 262}. He attempts to interpret, in the described manner, the equations of my continued studies concerning the

---

[1]   Footnotes refer to this translation. Boltzmann's citations to "this book", "earlier work" *etc.*, and citations in footnotes, have been replaced by explicit references indicated in the text by {}. Wiss. Abhand. refers to his Collected Works.

[2]   E is not energy; in later publications Boltzmann used the symbol H, as in the H-theorem.

equilibrium of heat particles. However, the line of reasoning of Mr. Meyer remained entirely unclear to me and I will return to my concerns with his approach on page 172 (Wiss. Abhand. Vol II).

We have to take here a totally different approach because it is our main purpose not to limit our discussion to thermal equilibrium, but to explore the relationship of this probabilistic formulation to the second theorem of the mechanical theory of heat. We want first to solve the problem which I referred to above and already defined in my "Remarks on some problems of the mechanical theory of heat" [3], namely to calculate the probability of state distributions from the number of different distributions. We want first to treat as simple a case as possible, namely a gas of rigid absolutely elastic spherical molecules trapped in a container with absolutely elastic walls. (Which interact with central forces only within a certain small distance, but not otherwise; the latter assumption, which includes the former as a special case, does not change the calculations in the least). Even in this case, the application of probability theory is not easy. The number of molecules is not infinite, in a mathematical sense, yet the number of velocities each molecule is capable of is effectively infinite. Given this last condition, the calculations are very difficult; to facilitate understanding, I will, as in earlier work, consider a limiting case.

## I. Kinetic Energy Has Discrete Values

We assume initially, each molecule is only capable of assuming a finite number of velocities, such as

$$0, \frac{1}{q}, \frac{2}{q}, \frac{3}{q}, \dots \frac{p}{q},$$

where $p$ and $q$ are arbitrary finite numbers. Upon colliding, two molecules may exchange velocities, but after the collision both molecules still have one of the above velocities, namely

$$0, \text{ or } \frac{1}{q}, \text{ or } \frac{2}{q}, \text{ etc. till } \frac{p}{q},$$

This assumption does not correspond to any realistic mechanical model, but it is easier to handle mathematically, and the actual problem to be solved is re-established by letting $p$ and $q$ go to infinity.

Even if, at first sight, this seems a very abstract way of treating the problem, it rapidly leads to the desired objective, and when you consider that in nature all infinities are but limiting cases, one assumes each molecule can behave in this fashion only in the limiting case where each molecule can assume more and more values of the velocity.

To continue, however, we will consider the kinetic energy, rather than the velocity of the molecules. Each molecule can have only a finite number of values for its kinetic energy. As a further simplification, we assume that the kinetic energies of each molecule form an arithmetic progression, such as the following:

$$0, \epsilon, 2\epsilon, 3\epsilon, \dots p\epsilon$$

---

[3]    Wiss Abhand, reprint 39.

We call $P$ the largest possible value of the kinetic energy, $p\epsilon$. Before impact, each of two colliding molecules shall have a kinetic energy of

$$0, \text{ or } \epsilon, \text{ or } 2\epsilon, \text{ } etc. \ldots p\epsilon$$

which means that after the collision, each molecule still has one of the above values of kinetic energy. The number of molecules in the vessel is $n$. If we know how many of these $n$ molecules have a kinetic energy of zero, how many have a kinetic energy of $\epsilon$ and so on, then we know the kinetic energy distribution. If at the beginning there is some state distribution among the gas molecules, this will in general be changed by the collisions. The laws governing this change have already been the subject of my previous investigations. But right away, I note that this is not my intention here; instead I want to establish the probability of a state distribution regardless of how it is created, or more specifically, I want to find all possible combinations of the $p+1$ kinetic energy values allowed to each of the $n$ molecules and then establish how many of these combinations correspond to each state distribution. The latter number then determines the likelihood of the relevant state distribution, as I have already stated in my published "Remarks about several problems in the mechanical theory of heat" (Wiss. Abhand. Vol II, reprint 39, p121.) As a preliminary, we will use a simpler schematic approach to the problem, instead of the exact case. Suppose we have $n$ molecules. Each of them is capable of having kinetic energy

$$0, \epsilon, 2\epsilon, 3\epsilon, \ldots p\epsilon.$$

and suppose these energies are distributed in all possible ways among the $n$ molecules, such that the total energy is a constant, e.g., $\lambda\epsilon = L$. Any such distribution, in which the first molecule may have a kinetic energy of e.g., $2\epsilon$, the second may have $6\epsilon$, and so on, up to the last molecule, we call a complexion, and so that each individual complexion can be easily enumerated, we write them in sequence (for convenience we divide through by $\epsilon$), specifying the kinetic energy of each molecule. We seek the number $\mathcal{P}$ of complexions where $w_0$ molecules have kinetic energy 0, $w_1$ molecules have kinetic energy $\epsilon$, $w_2$ have kinetic energy $2\epsilon$, up to the $w_p$ which have kinetic energy $p\epsilon$. We said, earlier, that given how many molecules have kinetic energy 0, how many have kinetic energy $\epsilon$, *etc.*, this distribution among the molecules specifies the number of $\mathcal{P}$ of complexions for that distribution; in other words, it determines the likelihood of that state distribution. Dividing the number $\mathcal{P}$ by the number of all possible complexions, we get the probability of the state distribution. Since a distribution of states does not determine kinetic energies exactly, the goal is to describe the state distribution by writing as many zeros as molecules with zero kinetic energy ($w_0$), $w_1$ ones for those with kinetic energy $\epsilon$ *etc.* All these zeros, ones, *etc.* are the elements defining the state distribution. It is now immediately clear that the number $\mathcal{P}$ for each state distribution is exactly the same as the number of permutations of which the elements of the state distribution are capable, and that is why the number $\mathcal{P}$ is the desired measure of the permutability of the corresponding distribution of states. Once we have specified every possible complexion, we have also all possible state distributions, the latter differing from the former only by immaterial permutations of molecular labels. All those complexions which contain the same number of zeros, the same number of ones *etc.*, differing from each other merely by different arrangements of elements, will result in the same state distribution; the number of complexions forming the same state distribution, and which we have denoted by $\mathcal{P}$, must be equal to the number of permutations which

the elements of the state distribution are capable of. In order to give a simple numerical example, take $n = 7, \lambda = 7, p = 7$, so $L = 7\epsilon, P = 7\epsilon$. With 7 molecules, there are 8 possible values for the kinetic energy, $0, \epsilon, 2\epsilon, 3\epsilon, 4\epsilon, 5\epsilon, 6\epsilon, 7\epsilon$ to distribute in any possible way such that the total kinetic energy $= 7\epsilon$. There are then 15 possible state distributions. We enumerate each of them in the above manner, producing the numbers listed in the second column of the following table of state distributions. The numbers in the first column label the different state distributions.

| | | $\mathcal{P}$ | | | $\mathcal{P}$ |
|---|---|---|---|---|---|
| 1. | 0000007 | 7 | 9. | 0001114 | 140 |
| 2. | 0000016 | 42 | 10. | 0001123 | 420 |
| 3. | 0000025 | 42 | 11. | 0001222 | 140 |
| 4. | 0000034 | 42 | 12. | 0011113 | 105 |
| 5. | 0000115 | 105 | 13. | 0011122 | 210 |
| 6. | 0000124 | 210 | 14. | 0111112 | 42 |
| 7. | 0000133 | 105 | 15. | 1111111 [1] | 1 |
| 8. | 0000223 | 105 | | | |

[1] The state distributions are so arranged that, read as a number, the rows are arranged in increasing order.

In the last column, under the heading $\mathcal{P}$ is the number of possible permutations of members for each state. The first state distribution, for example, has 6 molecules with zero kinetic energy, and the seventh has kinetic energy $7\epsilon$. So $w_0 = 6, w_7 = 1, w_2 = w_3 = w_4 = w_5 = w_6 = 0$. It is immaterial which molecule has kinetic energy $7\epsilon$. So there are 7 possible complexions which represent this state distribution. Denoting the sum of all possible complexions, 1716, by $J$ then the probability of the first state distribution is $7/J$; similarly, the probability of the second state distribution is $42/J$; the most probable state distribution is the tenth, as its elements permit the greatest number of permutations. Hereon, we call the number of permutations the relative likelihood of the state distribution; this can be defined in a different way, which we next illustrate with a specific numerical example, since generalization is straightforward. Suppose we have an urn containing an infinite number of paper slips. On each slip is one of the numbers 0, 1, 2, 3, 4, 5, 6, 7; each number is on the same amount of slips, and has the same probability of being picked. We now draw the first septet of slips, and note the numbers on them. This septet provides a sample state distribution with a kinetic energy of $\epsilon$ times the number written on the first slip for molecule 1, and so forth. We return the slips to the urn, and draw a second septet which gives us a second state distribution, *etc*. After we draw a very large number of septets, we reject all those for which the total does not equal 7. This still leaves a large number of septets. Since each number has the same probability of occurrence, and the same elements in a different order form different complexions, each possible complexion will occur equally often. By ordering the numbers within each septet by size, we can classify each into one of the fifteen cases tabulated above. So the number of septets which fall into the class 0000007 relative to the 0000016 class will be 7:42. Similarly for all the other septets. The most likely state distribution is the one which produces the most septets, namely the 10th.

(LB footnote: [4] If we divide the number of septets corresponding to a particular state by the total number of septets, we obtain the probability distribution. Instead of discarding all septets whose total is not 7 we could, after drawing a slip remove from the urn all those other slips for which a total of 7 is now impossible, e.g., on drawing a slip with 6 on it, all other slips except those with 0 or 1 would be removed. If the first 6 slips all had 0 on them, only slips with 7 on them would be left in the urn. One more thing should be noted at this point: We construct all possible complexions. If we denote by $\bar{w}_0$ the arithmetic mean of all values of $w_0$ which belong to the different complexions, and form analogous expressions $\bar{w}_1, \bar{w}_2 \ldots$, in the limit these quantities would also form the same state distribution.)

(Translators' note: Boltzmann's comments on the results of Mr. Oskar Meyer beginning "Ich will hier einige Worte über die von Hrn. Oskar Meyer..." on p172 (Wiss. Ab.) and ending with "...Bearbeitung des allgemeinen Problems zurückkehren." on p175 (Wiss. Ab.) are of historical interest only and are omitted.)

The first task is to determine the permutation number, previously designated by $\mathcal{P}$, for any state distribution. Denoting by $J$ the sum of the permutations $\mathcal{P}$ for all possible state distributions, the quotient $\mathcal{P}/J$ is the state distribution's probability, henceforth denoted by $W$. We would first like to calculate the permutations $\mathcal{P}$ for the state distribution characterized by $w_0$ molecules with kinetic energy 0, $w_1$ molecules with kinetic energy $\epsilon$, *etc.* It must be understood that

$$w_0 + w_1 + w_2 + \ldots + w_p = n \tag{1}$$

$$w_1 + 2w_2 + 3w_3 + \ldots + pw_p = \lambda \tag{2}$$

because the total number of molecules is $n$, and the total kinetic energy is $\lambda\epsilon = L$. Describing the state distribution as before, a complexion has $w_0$ molecules with zero energy, $w_1$ with one unit, and so on. The permutations, $\mathcal{P}$, arise since of the $n$ elements $w_0$ are mutually identical. Similarly with the $w_1, w_2$, *etc.* elements. The total number of permutations is well known

$$\mathcal{P} = \frac{n!}{w_0! w_1! \ldots} \tag{3}$$

The most likely state distribution will be for those $w_0, w_1 \ldots$ values for which $\mathcal{P}$ is a maximum or since the numerator is a constant, for which the denominator is a minimum. The values $w_0, w_1$ must simultaneously satisfy the two constraints (1) and (2). Since the denominator of $\mathcal{P}$ is a product, it is easiest to determine the minimum of its logarithm, that is the minimum of

$$M = \ln[(w_0!] + \ln[(w_1]! \ldots \tag{4}$$

here $\ln$ is the natural logarithm[5]. It is natural in our problem that only integer values of $w_0, w_1 \ldots$ are meaningful. However to apply differential calculus, we will allow non-integer values, and so find the minimum of the expression

$$M_1 = \ln\Gamma(w_0+1) + \ln\Gamma(w_1+1) + \ldots \tag{4a}$$

---

[4]   The paper contains two long footnotes of a scientific nature. For readability these are interpolated in the text where they are cited, as indicated by the text beginning "LB footnote:"

[5]   The ambiguous symbol "l" for logarithm in the original text has been replaced throughout by "ln".

which is identical to (4) for integer values of $w_0$, $w_1$ ..... We then get the non-integer values which for constraints (1) and (2) maximize $M_1$ [6]. The solution to the problem will in any case be obtained if for $w_0$, $w_1$, *etc.* we select the closest set of integer values. If here and there a deviation of a few integers is required, the nearest complexion is easily found. The minimum of $M_1$ is found by adding to both sides of the equation for $M_1$ Equation (1) multiplied by the constant $h$, and Equation (2) multiplied by the constant $k$, and setting the partial derivatives with respect to each of the variables $w_0$, $w_1$, $w_2$ ... to zero. We thus obtain the following equations:

$$\frac{d \ln \Gamma(w_0 + 1)}{dw_0} + h = 0,$$

$$\frac{d \ln \Gamma(w_1 + 1)}{dw_1} + h + k = 0,$$

$$\frac{d \ln \Gamma(w_2 + 1)}{dw_2} + h + 2k = 0$$

$$\vdots \quad \vdots \quad \vdots$$

$$\frac{d \ln \Gamma(w_p + 1)}{dw_p} + h + pk = 0,$$

which leads to

$$\frac{d \ln \Gamma(w_1 + 1)}{dw_1} - \frac{d \ln \Gamma(w_0 + 1)}{dw_0}$$

$$= \frac{d \ln \Gamma(w_2 + 1)}{dw_2} - \frac{d \ln \Gamma(w_1 + 1)}{dw_1}$$

$$= \frac{d \ln \Gamma(w_3 + 1)}{dw_3} - \frac{d \ln \Gamma(w_2 + 1)}{dw_2} \ldots \tag{5}$$

Exact solution of the problem through evaluation of the gamma function integral is very difficult; fortunately the general solution for arbitrary finite values of $p$ and $n$ does not interest us here, but only the solution for the limiting case of larger and larger number of molecules. Then the numbers $w_0$, $w_1$, $w_2$ *etc.* become larger and larger, so we introduce the function $\phi(x) = \ln \Gamma(x + 1) - x(\ln x - 1) - \frac{1}{2} \ln 2\pi$ [7]. Then we can write the first equation of (5) as follows

$$\ln w_1 + \frac{d\phi(w_1)}{dw_1} - \ln w_0 - \frac{d\phi(w_0)}{dw_0} = \ln w_2 + \frac{d\phi(w_2)}{dw_2} - \ln w_1 - \frac{d\phi(w_1)}{dw_1} \tag{6}$$

Similarly for the other equations of (5). It is also well known that

$$\phi(x) = -\frac{1}{2} \ln x + \frac{1}{12x} + \ldots \tag{6a}$$

This series is not valid for$v$ $x = 0$, but here $x!$ and $\sqrt{2\pi}(x/e)^x$ should have the same value, and $\phi(x) = 0$. Therefore the problem of finding the minimum of $w_0! w_1! w_2! ...$ is replaced by the easier problem of finding the minimum of

$$\sqrt{2\pi} \left(\frac{w_0}{e}\right)^{w_0} \sqrt{2\pi} \left(\frac{w_1}{e}\right)^{w_1} \sqrt{2\pi} \left(\frac{w_2}{e}\right)^{w_2} \ldots$$

---

[6]  The original text reads as "maximized" but should mean "minimized".

[7]  Boltzmann approximates $\ln x!$ by $x \ln x - x + \frac{1}{2} \ln (2\pi)$ rather than $(x + \frac{1}{2}) \ln x - x + \frac{1}{2} \ln (2\pi)$ as is now usual. For $x \gg 30$ the relative difference is small.

providing $w$ is not zero, even at moderately large values of $p$ and $n$ both problems have matching solutions. From line (6a) it follows

$$\frac{d\phi(w_0)}{dw_0} = -\frac{1}{2w_0} - \frac{1}{12w_0^2} - \cdots, \tag{6b}$$

which for larger and larger values of $w_0$ or $\ln w_0$ vanishes, the same also applies to the other $w's$, so the Equation (6) can be written as follows

$$\ln w_1 - \ln w_0 = \ln w_2 - \ln w_1$$

or

$$\frac{w_1}{w_0} = \frac{w_2}{w_1}$$

likewise the equations for the remaining $w's$ are

$$\frac{w_2}{w_1} = \frac{w_3}{w_2} = \frac{w_4}{w_3} = \ldots \tag{6c}$$

One sees immediately that by neglecting the expression (6b), the minimum of the denominator of

$$\frac{\sqrt{2\pi}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi}\left(\frac{w_0}{e}\right)^{w_0}\sqrt{2\pi}\left(\frac{w_1}{e}\right)^{w_1}\ldots}$$

is found instead of the minimum of the denominator of (3). So for problems involving $w!$, use of a well known approximation (See Schlömilch's Comp. S. 438) amounts to substitution of $\sqrt{2\pi}(w/e)^w$ for $w!$.

If we denote the common value of the quotient (6c) by $x$ we obtain

$$w_1 = w_0 x, w_2 = w_0 x^2, w_3 = w_0 x^3, \textit{ etc.} \tag{7}$$

The two Equations (1) and (2) become

$$w_0\left(1 + x + x^2 + \ldots + x^p\right) = n, \tag{8}$$

$$w_0\left(x + 2x^2 + 3x^3 + \ldots + px^p\right) = \lambda, \tag{9}$$

One sees immediately that these equations differ negligibly from Equation (42) and the preceding ones from my earlier work "Study of the thermal equilibrium of gas molecules".

We can use the last equation to write

$$w_0 \cdot \frac{x^{p+1} - 1}{x - 1} = n$$

$$w_0 x \cdot \frac{d}{dx}\left[\frac{x^{p+1} - 1}{x - 1}\right] = \lambda \tag{10}$$

Carrying out the differentiation in the last equation

$$w_0 x \frac{px^{p+1} - (p+1)x^p + 1}{(x-1)^2} = \lambda \tag{11}$$

Dividing this equation by Equation (10) gives

$$\frac{px^{p+2} - (p+1)x^{p+1} + x}{(x^{p+1} - 1)(x - 1)} = \frac{\lambda}{n}$$

or

$$(pn - \lambda)x^{p+2} - (pn + n - \lambda)x^{p+1} + (n + \lambda)x - \lambda = 0. \tag{12}$$

One can see immediately from Descartes' theorem [8] that this equation cannot have more than three real positive roots, of which two are $= 1$. Again it is easy to see that both roots are not solutions of Equations (8) and (9), and also do not solve the problem, but that they showed up in the final equation merely as a result of multiplying by the factor $(x - 1)^2$. To be convinced of this, one need only derive the final equation directly by dividing the Equations (8) and (9). Following this division and having removed the variable $x$ from the denominator and collecting powers of $x$ throughout, we get the equation

$$(np - 1)x^p + (np - n - \lambda)x^{p-1} + (np - 2n - \lambda)x^{p-2} + \ldots (n - \lambda)x - \lambda = 0, \tag{13}$$

which is an equation of $p^{th}$ degree, and whose roots supply the solution to the problem. Thus Equation (12) cannot have more positive roots than the solution requires. Negative or complex roots have no meaning for the solution to the problem. We note again that the largest allowed kinetic energy $P = p\epsilon$, is very large compared to the mean kinetic energy of a molecule

$$\frac{L}{n} = \frac{\lambda\epsilon}{n}$$

from which it follows that $p$ is very large compared to $\lambda/n$. The polynomial Equation (13), which shares the same real roots with Equation (12), is negative for $x = 0$, $x = 1$, however it has the value

$$n(p + 1)\left(\frac{p}{2} - \frac{\lambda}{n}\right),$$

which is positive and very large, since $p$ is very large compared to $n$. The only positive root occurs for $x$ between $0$ and $1$, and we obtain it from the more convenient Equation (12). Since $x$ is a proper fraction, then the $p^{th}$ [9] and $(p + 1)^{th}$ powers are smaller and can be neglected, in which case we obtain

$$x = \frac{\lambda}{n + \lambda}$$

This is the value to which $x$ tends for large $p$, and one can see the important fact that for reasonably large values of $p$ the value of $x$ depends almost exclusively on the ratio $\lambda/n$, and varies little with either $\lambda$ or $n$ providing their ratio is constant. Once one has found $x$, it follows from Equation (10) that

$$w_0 = \frac{1 - x}{1 - x^{p+1}}n \tag{14}$$

and Equation (7) gives the values of the remaining $w's$. It is seen from the quotients

$$\frac{w_0}{n}, \; \frac{w_1}{n}, \; \frac{w_2}{n}, \; etc.$$

---

[8]  Cardano's formula
[9]  This appears to be a typographic error. The $(p + 2)^{th}$ power makes mathematical sense.

that the probabilities of the various kinetic energy values for larger $p$ are again dependent almost exclusively on the mean energy of the molecule. For infinitely large $p$ we obtain the following limiting values

$$w_0 = \frac{n^2}{n + \lambda}, \quad w_1 = \frac{n^2\lambda}{(n + \lambda)^2}, \quad w_2 = \frac{n^2\lambda^2}{(n + \lambda)^3} \; etc. \qquad (15)$$

To establish whether we have a maximum or have a minimum, we need to examine the second variation of the Equation (4). We note that $w_0, w_1, w_2$, *etc.* are very large, so we can use the approximation formula for $\ln \Gamma(w + 1)$

$$w \cdot (\ln w - 1) - \frac{1}{2} \ln w - \frac{1}{2} \ln (2\pi) + \frac{1}{12w} + \; etc.$$

and neglecting terms which have second or higher powers of w in the denominator, obtain

$$\delta^2 M = \frac{(\delta w_0)^2}{w_0} + \frac{(\delta w_1)^2}{w_1} + \dots$$

so we do in fact have a minimum. I also want to remark on the size of the term previously designated $J$. One easily finds that $J$ is given by the following binomial coefficient

$$J = \binom{\lambda + n - 1}{\lambda};$$

when you neglect terms that diminish with increasing $\lambda$ or $n$

$$J = \frac{1}{\sqrt{2\pi}} \frac{(\lambda + n - 1)^{\lambda + n - \frac{1}{2}}}{(n - 1)^{n - \frac{1}{2}} \lambda^{\lambda + \frac{1}{2}}}$$

Now $\lambda\epsilon/n$ is equal to the average kinetic energy $\mu$ of a molecule, therefore

$$\frac{\lambda}{n} = \frac{\mu}{\epsilon},$$

So for large numbers one has

$$(\lambda + n - 1)^{\lambda + n - \frac{1}{2}} == \lambda^{\lambda + n - \frac{1}{2}} \left[ \left( 1 + \frac{n - 1}{\lambda} \right)^\lambda \right]^{1 + \frac{2n-1}{2\lambda}} = \lambda^{\lambda + n - \frac{1}{2}} \cdot e^{n-1}$$

Therefore

$$J = \frac{1}{\sqrt{2\pi}} \frac{\lambda^{n-1} e^{n-1}}{(n - 1)^{n - \frac{1}{2}}},$$

therefore, neglecting diminishing terms

$$\ln J = n \ln \frac{\lambda}{n} + n - \ln \lambda + \frac{1}{2} \ln n - 1 - \frac{1}{2} \ln (2\pi)$$

It goes without saying that these formulas are not derived here solely for finite $p$ and $n$ values, because these are unlikely to be of any practical importance, but rather to obtain formulas which provide the correct limiting values when $p$ and $n$ become infinite.

Nevertheless, it may help to demonstrate, with specific examples of only moderately large values of $p$ and $n$ that these formulas are quite accurate, and though approximate, are of some value even here.

We first consider the earlier example, where $n = \lambda = 7$, *i.e.*, the number of molecules is 7, and the total kinetic energy is $7\epsilon$, and so the mean kinetic energy is also $\epsilon$. Suppose first, that $p = 7$, so each molecule can only have $0, \epsilon, 2\epsilon, 3\epsilon, \ldots 7\epsilon$ of kinetic energy. Then Equation (12) becomes

$$6x^9 - 7x^8 + 2x - 1 = 0, \tag{16}$$

from which it follows

$$x = \frac{1}{2} + \frac{7}{2}x^8 - 3x^9. \tag{17}$$

Since $x$ is close to $\frac{1}{2}$, we can set $x = \frac{1}{2}$ in the last two very small terms on the right-hand side, and obtain

$$x = \frac{1}{2} + \frac{1}{2^9}(7 - 3) = \frac{1}{2} + \frac{1}{2^7} = 0.5078125$$

You could easily substitute this value for $x$ back into the right side of Equation (17) and obtain a better approximation for $x$; since we already have an approximate value for $x$, a more rapid approach is to apply the ordinary Newton iteration method to Equation (16) which results in

$$x = 0.5088742\ldots$$

From this, one finds in accordance with Equations (7)

$$
\begin{array}{ll}
w_0 = 3.4535 & w_4 = 0.2316 \\
w_1 = 1.7574 & w_5 = 0.1178 \\
w_2 = 0.8943 & w_6 = 0.0599 \\
w_3 = 0.4551 & w_7 = 0.0304
\end{array}
$$

These numbers satisfy the condition that

$$\sqrt{2\pi}\left(\frac{w_0}{e}\right)^{w_0} \cdot \sqrt{2\pi}\left(\frac{w_1}{e}\right)^{w_1} \ldots \textit{etc.}$$

is minimized, while the minimized variables w obey the two constraints

$$w_0 + w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 = 7,$$

$$w_1 + 2w_2 + 3w_3 + 4w_4 + 5w_5 + 6w_6 + 7w_7 = 7, \tag{18}$$

which minimum, incidentally because of the first of Equations (18) coincides with the minimum of $(w_0)^{w_0}(w_1)^{w_1}\ldots$ This provides only an approximate solution to our problem, which asks for so many $(w_0)$ zeros, so many $(w_1)$ ones *etc.* with as many permutations as the resulting complexion permits, while the $w's$ simultaneously satisfy the constraints (18). Since $p$ and $n$ here are very small, one hardly expects any great accuracy, yet you already get the solution to the permutation problem by taking the nearest integer for each $w$, with the exception of $w_3$, for which you have to assign the value of 1 instead of 0,4551. In this manner it is apparent

$$w_0 = 3, w_1 = 2, w_2 = w_3 = 1, w_4 = w_5 = w_6 = w_7 = 0$$

and in fact we saw in the previous table that the complexion of 0001123 has the most permutations. We now consider the same special case with $n = \lambda = 7$, but set $p = \infty$; that is, the molecules may have

kinetic energies of $0, 1, 2, 3 \ldots \infty$. We know then that the values of the variables $w$ will vary little from those of the former case. In fact we obtain

$$x = \frac{1}{2}; w_0 = \frac{7}{2} = 3.5, \;\; w_1 = \frac{w_0}{2} = 1.75, \;\; w_2 = \frac{w_1}{2} = 0.875 \; etc.$$

We consider a little more complicated example. Take $n = 13, \lambda = 19$, but we only treat the simpler case where $p = \infty$. Then we have

$$
\begin{aligned}
x &= 19/32 & w_4 &= 0.65605 \\
w_0 &= 5.28125 & w_5 &= 0.38950 \\
w_1 &= 3.13574 & w_6 &= 0.23133 \\
w_2 &= 1.86815 & &\vdots \\
w_3 &= 1.10493 & &\vdots
\end{aligned}
$$

Substituting here for the $w's$ the nearest integers, we obtain

$$w_0 = 5, w_1 = 3, w_2 = 2, w_3 = 1, w_4 = 1, w_5 = w_6 \ldots 0,$$

Already from the fact that $w_0 + w_1 + \ldots$ should $= 13$, it is seen that again one of the $w's$ must be increased by one unit. From those $w's$ that are set $= 0$, $w_5$ differs least from the next highest integer. We want therefore $w_5 = 1$, and obtain the complexion

$$0000011122345$$

whose digit sum is in fact $= 19$. The number of permutations this complexion is capable of is

$$\frac{13!}{5! \, 3! \, 2!} = \frac{13!}{4! \, 3! \, 2!} \cdot \frac{1}{5}$$

A complexion whose sum of digits is also $= 19$, and which one might suppose is capable of very many permutations, would be the following:

$$0000111222334.$$

The number of permutations is

$$\frac{13!}{4! \, 3! \, 3! \, 2!} = \frac{13!}{4! \, 3! \, 2!} \cdot \frac{1}{6}$$

This is less than the number of permutations of the first complexion we found from the approximate formula. Likewise, we expect that the number of permutations of the two complexions

$$0000111122335$$

and

$$0000111122344$$

is smaller still. This is, for both complexions

$$\frac{13!}{4! \, 4! \, 2! \, 2!} = \frac{13!}{4! \, 3! \, 2!} \cdot \frac{1}{8}$$

Other possible complexions are capable of still less permutations, and it would be quite superfluous to follow these up here. It is seen from the examples given here that the above formula, even for very small values of $p$ and $n$ gives values of $w$ within one or two units of the true values. In the mechanical theory of heat we are always dealing with extremely large numbers of molecules, so such small differences disappear, and our approximate formula provides an exact solution to the problem. We see also that the most likely state distribution is consistent with that known from gases in thermal equilibrium. According to Equation (15) the probability of having a kinetic energy $s\epsilon$ is given by

$$w_s = \frac{n^2}{n+\lambda} \cdot \left(\frac{\lambda}{n+\lambda}\right)^s$$

since $\lambda\epsilon/n$ is equal to the average kinetic of a molecule $\mu$, which is finite, so $n$ is very small compared to $\lambda$. So the following approximations

$$\frac{n^2}{n+\lambda} = \frac{n^2}{\lambda} = \frac{n\epsilon}{\mu}, \frac{\lambda}{n+\lambda} = 1 - \frac{n}{\lambda} = e^{-\frac{n}{\lambda}} = e^{-\frac{\epsilon}{\mu}}$$

hold. From which it follows that

$$w_s = \frac{n\epsilon}{\mu} e^{-\frac{\epsilon s}{\mu}}.$$

To achieve a mechanical theory of heat, these formulas must be developed further, particularly through the introduction of differentials and some additional considerations.

## II. Kinetic Energies Exchange in a Continuous Manner

In order to introduce differentials into our formula we wish to illustrate the problem in the same manner as indicated on p171 (Wiss. Abhand. vol II) because this seems to be the best way to clarify the matter. Here each molecule was only able to have one of $0, \epsilon, 2\epsilon, \ldots p\epsilon$ values for kinetic energy. We generated all possible complexions, *i.e.*, all the ways of distributing $1 + p$ values of the kinetic energy among the molecules, yet subject to the constraints of the problem, using a hypothetical urn containing infinitely many paper slips. Equal numbers of paper slips have kinetic energy values $0, \epsilon$ *etc.* written on them. To generate the first complexion, we draw a slip of paper for each molecule, and note the value of the kinetic energy assigned in this way to each molecule. Very many complexions are generated in the same way, they are assigned to this or that state distribution, and then we determine the most probably one. That state distribution which has the most complexions, we consider as the most likely, or corresponding to thermal equilibrium. Proceeding to the continuous kinetic energy case the most natural approach is as follows:

Taking $\epsilon$ to be some very small value, we assume that in the urn are very many slips of paper labeled with kinetic energy values between $0$ and $\epsilon$. In the urn are also equal numbers of paper slips labeled with kinetic energy values between $\epsilon$ and $2\epsilon$, $2\epsilon$ and $3\epsilon$ up to infinity. Since $\epsilon$ is very small, we can regard all molecules with kinetic energy between $x$ and $x + \epsilon$ as having the same kinetic energy. The rest of the calculation proceeds as in Section I above. We assume some complexion has been drawn; $w_0$ molecules have kinetic energy between $0$ and $\epsilon$, $w_1$ molecules have values between $\epsilon$ and $2\epsilon$, $w_2$ have values between $2\epsilon$ and $3\epsilon$, *etc.*

Here, because the variables $w_0, w_1, w_2$ *etc.* will be infinitely small, of the order of magnitude of $\epsilon$, we prefer to write them as

$$w_0 = \epsilon f(0), w_1 = \epsilon f(\epsilon), w_2 = \epsilon f(2\epsilon) \text{ etc.} \tag{19}$$

The probability of the state distribution in question is given, exactly as in Section I, by the number of permutations that the elements of the state distribution are capable of, e.g., by the number

$$\frac{n!}{w_0! w_1! w_2! \ldots}$$

Again, the most likely state distribution, which corresponds to thermal equilibrium, is defined by the maximum of this expression, that is, when the denominator is minimized. We use again the reasonable approximation of Section I, replacing $w!$ by the expression

$$\sqrt{2\pi} \left(\frac{w}{e}\right)^w$$

We can omit the term $\sqrt{2\pi}$ since it is a constant factor in the minimization; The key again is to replace minimization of the denominator with minimization of its logarithm; then we obtain the condition for thermal equilibrium, that

$$M = w_0 \ln w_0 + w_1 \ln w_1 + w_2 \ln w_2 + \ldots - n$$

is a minimum; while again satisfying the two constraints

$$n = w_0 + w_1 + w_2 + \ldots, \tag{20}$$

$$L = \epsilon w_1 + 2\epsilon w_2 + 3\epsilon w_3 + \ldots, \tag{21}$$

which are identical with Equations (1) and (2) of Section I. Using Equation (19) here, we replace the variables $w$ by the function $f$ and obtain thereby

$$M = \epsilon[f(0) \ln f(0) + f(\epsilon) \ln f(\epsilon) + f(2\epsilon) \ln f(2\epsilon) + \ldots] + \epsilon \ln \epsilon[f(0) + f(\epsilon) + f(2\epsilon) + \ldots] - n$$

and Equations (20) and (21) become

$$n = \epsilon[f(0) + f(\epsilon) + f(2\epsilon) + \ldots], \tag{22}$$

$$L = \epsilon[\epsilon f(\epsilon) + 2\epsilon f(2\epsilon) + 3\epsilon f(3\epsilon) \ldots]. \tag{23}$$

Using Equation (22) the expression for M can also be written as

$$M = \epsilon[f(0) \ln f(0) + f(\epsilon) \ln f(\epsilon) + f(2\epsilon) \ln f(2\epsilon) + \ldots] - n + n \ln \epsilon$$

Since $n$ and $\epsilon$ are constant, (because $\epsilon$ has the same value for all possible complexions, and is constant between different state distribution), one can minimize

$$M' = \epsilon[f(0) \ln f(0) + f(\epsilon) \ln f(\epsilon) + f(2\epsilon) \ln f(2\epsilon) + \ldots] \tag{24}$$

instead. As $\epsilon$ is made still smaller, the allowed values of kinetic energy approach a continuum. For vanishingly small $\epsilon$, various sums in Equations (22)–(24) can be written in the form of integrals, leading to the following equations

$$M' = \int_0^\infty f(x) \ln f(x) dx, \tag{25}$$

$$n = \int_0^\infty f(x) dx, \tag{26}$$

$$L = \int_0^\infty x f(x) dx, \tag{27}$$

The functional form of $f(x)$ is sought which minimizes expression (25) subject to the constraints (26) and (27), so one proceeds as follows: To the right side of Equation (25) one adds Equation (26) multiplied by a constant k, and Equation (27) multiplied by a constant h. The resulting integral is

$$\int_0^\infty [f(x) \ln f(x) + k f(x) + h x f(x)] dx$$

where $x$ is the independent variable, and $f$ is the function to be varied. This results in

$$\int_0^\infty [\ln f(x) + k + hx] \delta f(x) dx$$

Setting the quantity which has been multiplied by $\delta f(x)$ in square brackets $= 0$, and solving for the function $f(x)$, we obtain

$$f(x) = Ce^{-hx} \tag{28}$$

here the constant $e^{-k-1}$ is denoted by C for brevity. The second variation of $M'$

$$\delta^2 M' = \int_0^\infty \frac{[\delta f(x)]^2}{f(x)} \cdot dx,$$

is necessarily positive, since $f(x)$ is positive for all values of $x$ lying between 0 and $\infty$. By the calculus of variations $M'$ is a minimum. From Equation (28), the probability that the kinetic energy of a molecule lies between $x$ and $x + dx$ at thermal equilibrium is

$$f(x) dx = Ce^{-hx} dx$$

The probability that the velocity of a molecule lies between $\omega$ and $\omega + d\omega$ would be

$$Ce^{-\frac{hm\omega^2}{2}} \cdot m\omega d\omega, \tag{29}$$

where $m$ is the mass of a molecule. Equation (29) gives the correct state distribution for elastic disks moving in two dimensions, for elastic cylinders with parallel axis moving in space, but not for elastic spheres which move in space. For the latter the exponential function must be multiplied by $\omega^2 d\omega$ not $\omega d\omega$. To get the right state distribution for the latter case we must set up the initial distribution of paper slips in our urn in a different way. To this point we assumed that the number of paper slips labeled with kinetic energy values between 0 and $\epsilon$ is the same as those between $\epsilon$ and $2\epsilon$. As also for slips with kinetic energies between $2\epsilon$ and $3\epsilon$, $3\epsilon$ and $4\epsilon$, *etc.*

Now, however, let us assume that the three velocity components along the three coordinate axes, rather than the kinetic energies, are written on the paper slips in the urn. The idea is the same: There are the same number of slips with $u$ between 0 and $\epsilon$, $v$ between 0 and $\zeta$, and $w$ between 0 and $\eta$. The number of slips with $u$ between $\epsilon$ and $2\epsilon$, $v$ between 0 and $\zeta$, and $w$ between 0 and $\eta$ is the same. Similarly, the number for which $u$ is between $\epsilon$ and $2\epsilon$, $v$ is between $\zeta$ and $2\zeta$, $w$ is between 0 and $\eta$. Generally, the number of slips for which $u, v, w$ are between the limits $u$ and $u + \epsilon$, $v$ and $v + \zeta$, $w$ and $w + \eta$ are the same. Here $u, v, w$ have any magnitude, while $\epsilon, \zeta, \eta$ are infinitesimal constants. With this one modification of the problem, we end up with the actual state distribution established in gas molecules.

(LB footnote: We can of course, instead of using finite quantities $\epsilon, \zeta, \eta$ and then taking the limit as they go to zero, write $du, dv, dw$ from the outset, then the distribution of paper slips in the urn must be such that the number for which $u, v, w$ are between $u$ and $u + du$, $v$ and $v + dv$, $w$ and $w + dw$ are proportional to the product $du\,dv\,dw$ and independent of $u, v,$ and $w$. The earlier distribution of slips in the urn is characterized by the fact that although $\epsilon$ could be replaced by $dx$, kinetic energies between 0 and $dx$, $dx$ and $2dx$, $2dx$ and $3dx$ *etc.* occurred on the same number of slips.)

If we now define

$$w_{abc} = \epsilon\zeta\eta f(a\epsilon, b\zeta, c\eta) \tag{30}$$

as the number of molecules of any complexion for which the velocity components lie between the limits $a\epsilon$ and $(a + 1)\epsilon$, $b\zeta$ and $(b + 1)\zeta$, $c\eta$ and $(c + 1)\eta$, the number of permutations, or complexions of these elements for any state distribution, becomes

$$\mathcal{P} = \frac{n!}{\prod\limits_{a=-p}^{a=+p}\prod\limits_{b=-q}^{b=+q}\prod\limits_{c=-r}^{c=+r} w_{abc}!} \tag{31}$$

where we first assume $u$ adopts only values between $-p\epsilon$ and $+p\epsilon$, $v$ between $-q\zeta$ and $+q\zeta$, $w$ between $-r\eta$ and $+r\eta$. Where again, the most likely state distribution occurs when this expression, or if you will, its logarithm, is maximum. We again substitute

$$n! \text{ by } \sqrt{2\pi}\left(\frac{n}{e}\right)^n \text{ and } w! \text{ by } \sqrt{2\pi}\left(\frac{w}{e}\right)^w,$$

where you can again immediately omit the factors of $\sqrt{2\pi}$ as they simply contribute additive constants $-\frac{1}{2}\ln 2\pi$ to $\ln\mathcal{P}$; omitting also the constant $n\ln n$ term, the requirement for the most probable state distribution is that the sum

$$-\sum_{a=-p}^{a=+p}\sum_{b=-q}^{b=+q}\sum_{c=-r}^{c=+r} w_{abc}\ln w_{abc}$$

is a maximum, which only differs from $\ln\mathcal{P}$ by an additive constant. The constraints that the number of molecules$= n$ and that the total kinetic energy$= L$, take the following form

$$n = \sum_{a=-p}^{a=p}\sum_{b=-q}^{b=q}\sum_{c=-r}^{c=r} w_{abc} \tag{32}$$

$$L = \frac{m}{2}\cdot\sum_{a=-p}^{a=p}\sum_{b=-q}^{b=q}\sum_{c=-r}^{c=r}(a^2\epsilon^2 + b^2\zeta^2 + c^2\eta^2)w_{abc} \tag{33}$$

Substituting for $w_{abc}$ using Equation (30), one immediately sees that the triple sums can in the limit be expressed as definite integrals; omitting an additive constant, the quantity to be maximized becomes

$$\Omega = - \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} f(u,v,w) \ln f(u,v,w) du dv dw, \tag{34}$$

The two constraint equations become

$$n = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} f(u,v,w) du dv dw \tag{35}$$

$$L = \frac{m}{2} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} (u^2 + v^2 + w^2) f(u,v,w) du dv dw \tag{36}$$

The variable $\Omega$, which differs from the logarithm of the number of permutations only by an additive constant, is of special importance for this work and we call it the permutability measure. I note, incidentally, that suppression of the additive constants has the advantage that the total permutability measure of two bodies is equal to the sum of the permutability measures of each body.

Thus it is the maximum of the quantity (34) subject to the constraints (35) and (36) that is sought. No further explanation of this problem is needed here; it is a special case of the problem I have already discussed in my treatise "On the thermal equilibrium of gases on which external forces act" [10] in the section which immediately precedes the appendix. There I provided evidence that this state distribution corresponds to the condition of thermal equilibrium. Thus, one is justified in saying that the most likely state distribution corresponds with the condition of thermal equilibrium. For if an urn is filled with slips of paper labeled in the manner described earlier, the most likely sampling will correspond to the state distribution for thermal equilibrium. We should not takes this for granted, however, without first defining what is meant by the most likely state distribution. For example, if the urn was filled with slips labeled in the original manner then the statement would be incorrect.

The reasoning needed to arrive at the correct state distribution will not escape those experienced in working with such problems. The same considerations apply to the following circumstance: If we group all the molecules whose coordinates at a particular time lie between the limits

$$\xi \text{ and } \xi + d\xi, \eta \text{ and } \eta + d\eta, \zeta \text{ and } \zeta + d\zeta, \tag{37}$$

and whose velocity components lie between the limits

$$u \text{ and } u + du, v \text{ and } v + dv, w \text{ and } w + dw \tag{38}$$

and let these molecules collide with other molecules under specific conditions, after a certain time their coordinates will lie between the limits

$$\Xi \text{ and } \Xi + d\Xi, H \text{ and } H + dH, Z \text{ and } Z + dZ, \tag{39}$$

---

[10] Wien. Ber. (1875) 72:427-457 (Wiss. Abhand. Vol. II, reprint 32).

and their velocity components will lie between the limits

$$U \text{ and } U + dU, V \text{ and } V + dV, W \text{ and } W + dW \tag{40}$$

Then at any time

$$d\xi \cdot d\eta \cdot d\zeta \cdot du \cdot dv \cdot dw = d\Xi \cdot dH \cdot dZ \cdot dU \cdot dV \cdot dW. \tag{41}$$

This is a general result. If at time zero the coordinates and velocity components of arbitrary molecules (material points) lie between the limits (37) and (38), and unspecified forces act between these molecules, so that at time $t$ the coordinates and velocity components lie between the limits (39) and (40), then Equation (41) is still satisfied.

(Full details, a rigorous treatment and an even more general theorem can be found in the book by Watson "A treatise on the kinetic theory of gases", p12)

If, instead of the velocity components, one uses the kinetic energy $x$ and the velocity direction defined by the two angles $\alpha$ and $\beta$, to describe the action of the forces, these variables would initially lie between the limits

$$\xi \text{ and } \xi + d\xi, \eta \text{ and } \eta + d\eta, \zeta \text{ and } \zeta + d\zeta,$$

$$x \text{ and } x + dx, \alpha \text{ and } \alpha + d\alpha, \beta \text{ and } \beta + d\beta,$$

and then after the action of the forces lie between the limits

$$\Xi \text{ and } \Xi + d\Xi, H \text{ and } H + dH, Z \text{ and } Z + dZ,$$

$$X \text{ and } X + dX, A \text{ and } A + dA, B \text{ and } B + dB,$$

and so

$$d\xi \cdot d\eta \cdot d\zeta \cdot \sqrt{x} \cdot dx \cdot \varphi(\alpha, \beta)d\alpha \cdot d\beta = d\Xi \cdot dH \cdot dZ \cdot \sqrt{X} \cdot dX \cdot \varphi(A, B)dA \cdot dB$$

So the product of the differentials $du \cdot dv \cdot dw$ becomes $dU \cdot dV \cdot dW$. Therefore the list of slips in the urn must be labeled uniformly with velocity components lying between $u$ and $u + du$, $v$ and $v + dv$, $w$ and $w + dw$, whatever values $u, v, w$ have. Given a certain value of the coordinates, the velocities must be described by the corresponding "moments". On the other hand $\sqrt{x}dx$ goes over to $\sqrt{X}dX$. With the introduction of kinetic energy, slips must be labeled so that you have the same number with kinetic energy between $x$ and $x + \sqrt{x}dx$ where $dx$ is constant but $x$ is completely arbitrary. This last sentence is in agreement with my "Remarks on some problems of the mechanical theory of heat" (Wiss. Abhand. Vol II, reprint 39 p121), where I demonstrated that this is the only valid way to find the most likely state distribution corresponding to the actual thermal equilibrium; here we have demonstrated *a posteriori* that this leads to the correct state distribution for thermal equilibrium, that which is the most likely in our sense.

Of course it is easy to analyze those cases where other conditions exist besides the principle of conservation of kinetic energy. Suppose for example, a very large number of molecules for whom (1) The total kinetic energy is constant; (2) The net velocity of the center of gravity in the directions of the $x$-axis; (3) $y$-axis and (4) $z$-axis are given. The question arises, what is the most probable distribution

of the velocity components among the molecules, using the term in the previous sense. We then have exactly the same problem, except with four constraints instead of one. The solution gives us the most probable state distribution

$$f(u, v, w) = Ce^{-h[(u-\alpha)^2+(v-\beta)^2+(w-\gamma)^2]},$$

where C, h, $\alpha, \beta, \gamma$ are constants. This is in fact the state distribution for a gas at thermal equilibrium at a certain temperature, not at rest, but moving with a constant net velocity. You can treat similar problems such as the rotation of a gas in the same manner, by adding in the appropriate constraint equations; which I have discussed in my essay "On the definition and integration of the equations of molecular motion in gases." (Wiss. Abhand. Vol II, reprint 36)

   Some comment regarding the derivation of Equation (34) from Equation (31) is required here. The formula for $x!$ is

$$\sqrt{2\pi x}\left(\frac{x}{e}\right)^x e^{\frac{1}{12x}+\cdots}.$$

The substitution of this into Equation (31) gives

$$\mathcal{P} = \frac{\sqrt{2\pi}n^{n+\frac{1}{2}} \cdot e^{\frac{1}{12n}+\cdots}}{(2\pi)^{p+q+r+\frac{3}{2}} \prod\limits_{a=-p}^{a=+p}\prod\limits_{b=-q}^{b=+q}\prod\limits_{c=-r}^{c=+r} (w_{abc})^{w_{abc}+\frac{1}{2}} \cdot e^{\frac{1}{12w_{abc}}+\cdots}}$$

from which it follows

$$\ln\mathcal{P} = \left(n+\frac{1}{2}\right)\ln n+\frac{1}{12n}+\ldots-(p+q+r+1)\ln 2\pi-\sum_{a=-p}^{a=+p}\sum_{b=-q}^{b=+q}\sum_{c=-r}^{c=+r}\left[\left(w_{abc}+\frac{1}{2}\right)\ln w_{abc}+\frac{1}{12w_{abc}}+\ldots\right]$$

First note, that in determining the magnitude of $\mathcal{P}$, in the limit of a very large number of molecules, $n$, (and thus also of $w_{abc}$), other small quantities such as $\epsilon, \zeta, \eta$ can be treated as infinitesimals. So all terms which have $n$ or $w_{abc}$ in the denominator can be neglected, and also the $\frac{1}{2}$ in the term $w_{abc}+\frac{1}{2}$ . The terms containing $w_{abc}$ scale with the total mass of the gas, while the related $\frac{1}{2}$ terms refer only to a single molecule. So the latter quantities can be neglected as the number of molecules increases. We then get

$$\ln\mathcal{P} = n\ln n - (p+q+r+1)\ln 2\pi - \sum_{a=-p}^{a=+p}\sum_{b=-q}^{b=+q}\sum_{c=-r}^{c=+r} w_{abc}\ln w_{abc}.$$

Substituting $\epsilon\zeta\eta f(a\epsilon, b\zeta, c\eta)$ , for $w_{abc}$ we obtain

$$\ln\mathcal{P} = n\ln n - (p+q+r+1)\ln 2\pi - n\ln(\epsilon\zeta\eta) - \sum_{a=-p}^{a=+p}\sum_{b=-q}^{b=+q}\sum_{c=-r}^{c=+r} \epsilon\zeta\eta f(a\epsilon, b\zeta, c\eta)\ln f(a\epsilon, b\zeta, c\eta).$$

One sees that aside from the triple sum, the terms on the right hand side are constant, and so can be omitted. We also let $\epsilon, \zeta, \eta$ decrease while $p, q, r$ increases infinitely, so the triple sum goes over into a triple integral over limits $-\infty$ to $+\infty$ and from $\ln\mathcal{P}$ we arrive immediately at the expression given by Equation (34) for the permutability measure $\Omega$. The critical condition is that the number of molecules is very large; this means that $w_{abc}$ is large compared to $\frac{1}{2}$; also that the velocity components between

the limits $a\epsilon$ and $(a+1)\epsilon$, $b\zeta$ and $(b+1)\zeta$, $c\eta$ and $(c+1)\eta$ are identical to those between the limits $u$ and $u + du$, $v$ and $v + dv$, $w$ and $w + dw$. This may appear strange at first sight, since the number of gas molecules is finite albeit large, whereas $du, dv, dw$ are mathematical differentials. But on closer deliberation this assumption is self-evident. For all applications of differential calculus to the theory of gases are based on the same assumption, namely: diffusion, internal friction, heat conduction, *etc.* In each infinitesimal volume element $dxdydz$ there are still infinitely many gas molecules whose velocity components lie between the limits $u$ and $u + du$, $v$ and $v + dv$, $w$ and $w + dw$. The above assumption is nothing more than that very many molecules have velocity components lying within these limits for every $u, v, w$.

### III. Consideration of Polyatomic Gas Molecules and External Forces

I will now generalize the formulas obtained so far, by first extending them to so-called polyatomic gas molecules and then including external forces and thereby finally beginning to extend the discussion to any solid and liquid. In order not to consider too many examples, I will in each case deal with the most important case, where, aside from the equation for kinetic energy, there is no other constraint.

The first generalization can be applied to our formulas without difficulty. So far, we assumed each molecule was an elastic sphere or a material point, so that its position in space was entirely defined by three variables (e.g., three orthogonal coordinates). We know that this is not the case with real gas molecules. We shall therefore assume that three coordinates are insufficient to completely specify the position of all parts of a molecule in space; rather $r$ variables will be necessary

$$p_1, p_2, p_3 \ldots p_r, \tag{42}$$

the so called generalized coordinates. Three of them, $p_1, p_2, p_3$ are the orthogonal coordinates of the center of mass of the molecule, the others can be either the coordinates of the individual atoms relative to the center of mass, the angular direction, or whatever specifies the location of every part of the molecule. We will also remove the restriction that only one type of gas molecule is present. We assume instead, there exists a second type whose every molecule has the generalized coordinates

$$p'_1, p'_2, p'_3 \ldots p'_{r'},$$

for the third type the generalized coordinates are

$$p''_1, p''_2, p''_3 \ldots p''_{r''},$$

If there are $v + 1$ types of molecules, the generalized coordinates of the final type are

$$p_1^{(v)}, p_2^{(v)}, p_3^{(v)} \ldots p_{r^{(v)}}^{(v)}.$$

The first three coordinates are always the orthogonal coordinates of the center of mass. Of course the necessary assumption is that many molecules of each type are present. Let $l$ be the total kinetic energy

of the first type of gas; $\chi$ is its potential energy [11] (so that $\chi + l$ is constant if internal forces only are acting). Furthermore

$$q_1, q_2, q_3 \ldots q_r$$

are the momentum coordinates corresponding to $p_1, p_2 \ldots p_r$. We can think of $l$ in terms of the coordinates $p_1, p_2 \ldots p_r$ and their derivatives with respect to time

$$\dot{p}_1, \dot{p}_2, \dot{p}_3 \ldots \dot{p}_r,$$

and denote the quantities $c_1(dl/d\dot{p}_1), c_2(dl/d\dot{p}_2)$ by $q_1, q_2 \ldots$, where $c_1, c_2 \ldots$ are arbitrary constants.

I would like to note here that in my essay "Remarks on some problems in the mechanical theory of heat" Section III (Wiss. Abhand. Vol II, reprint 39) the indexed variables designated $p_i$ referred to coordinate derivatives, while here they have been designated by $q$; this mistake would probably not have caused any misunderstanding.

We denote with the appropriate accents the analogous quantities for other types of molecules. According to the calculations of Maxwell, Watson and myself, in a state of thermal equilibrium, the number of molecules for which the magnitudes of $p_4, p_5 \ldots p_r, q_1, q_2, q_3 \ldots q_r$ lie between the limits

$$p_4 \text{ and } p_4 + dp_4, p_5 \text{ and } p_5 + dp_5 \ etc. q_r \text{ and } q_r + dq_r \tag{43}$$

is given by

$$Ce^{-h(\chi+l)}dp_4 dp_5 \ldots dq_r \tag{44}$$

where $C$ and $h$ are constants, independent of $p$ and $q$. Analogous expressions hold of course for the other molecular species with the same value of $h$, but differing values of $C$. Exactly the same equation as (44) is also obtained using the methods of Sections I and II. Consider all those molecules of the first type, for which the variables $p_4, p_5 \ldots q_r$ at some time 0 lie between the limits (43), after a lapse of some time $t$, the values of the same variables lie between the limits

$$P_4 \text{ and } P_4 + dP_4, P_5 \text{ and } P_5 + dP_5 \ etc. Q_r \text{ and } Q_r + dQ_r \tag{45}$$

The general principle already invoked, gives the following equation

$$dp_4 \cdot dp_5 \ldots dq_r = dP_4 \cdot dP_5 \ldots dQ_r, \tag{46}$$

There is of course also

$$dp_1 \cdot dp_2 \cdot dp_3 = dP_1 \cdot dP_2 \cdot dP_3.$$

So that, in fact, for the variables

$$p_4, p_5, \ldots q_r$$

the product of their differentials does not change during a constant time interval. So we must now imagine $v + 1$ urns. In the first are slips of paper, upon which are written all possible values of the

---

[11] The quantity $\chi$ called "Kraftfunktion" or "Ergal" by Boltzmann is translated as potential energy.

variables $p_4, p_5 \ldots q_r$; and the number of slips which have values within the limits of Equation (43) is such that, when divided by the product $dp_4 \cdot dp_5 \ldots q_r$ it is a constant.

Similarly for the labeling of the slips with the variables $p'_4, p'_5 \ldots q'_r$ in the second urn, except that for the latter the constant can have a different value. The same applies to the other urns. We draw from the first urn for each molecule of the first type, from the second urn for molecules of the second type, *etc.* We now suppose that the values of the variables for each molecule are determined by the relevant drawings. It is of course entirely chance that determines the state distributions for the gas molecules, and we must first discard those state distributions which do not have the prescribed value for total kinetic energy. It will then be most likely that the state distribution described by (44) will be drawn, *i.e.*, that one corresponding to thermal equilibrium. The proof of this is straightforward. So the results found in the first two sections can be readily generalized to this case.

We want to generalize the problem further, assuming that the gas is composed of molecules specified exactly as before. But now so called external forces are acting, e.g., those like gravity which originate outside the gas. For details on the nature of these external forces, and how to treat them, see my treatise "On the thermal equilibrium of gases on which external forces act" [12]. The essence of the solution to the problem remains the same. Only now the state distribution will no longer be the same at all points of the vessel containing the gas; therefore $dp_1 \cdot dp_2 \cdot dp_3 = dP_1 \cdot dP_2 \cdot dP_3$ will no longer hold. We will now understand the generalized coordinates $p_1, p_2 \ldots p_r$ more generally to determine the absolute position of the molecule in space and the relative position of its constituents.

The notion that $p_1, p_2, p_3$ are just the orthogonal coordinates of the center of gravity is dropped. The same is true for the molecules of all the other types of gas. There is one further point to notice. Previously the only necessary condition was that throughout the vessel very many molecules of each type were present; now it is required that even in a small element of space, over which the external forces do not vary significantly in either size or direction, very many molecules are present (a condition, incidentally, which must hold for any theoretical treatment of problems where external forces on gases come into play). This is because our method of sampling presupposes that the states of many molecules can be considered equivalent, in the sense that the state distribution is not changed when the states of these molecules are exchanged. The probability of a state distribution is then determined by the number of complexions of which this state distribution is capable of.

This is why, for the case just considered, with $v + 1$ molecular species present, $v + 1$ urns must be constructed.

We assume first, that a complexion has been drawn where

$$w_{000\ldots} = f(0, 0, 0 \ldots) \alpha \beta \gamma \ldots$$

molecules whose variables $p_1, p_2 \ldots q_r$ lie between limits $0$ and $\alpha$, $0$ and $\beta$, $0$ and $\gamma$ *etc.* Furthermore, exactly

$$w_{10000\ldots} = f(\alpha, 0, 0 \ldots) \alpha \beta \gamma \ldots$$

---

[12]  Wien. Ber. (1875) 72:427-457.

molecules have the same variables within the limits $\alpha$ and $2\alpha$, 0 and $\beta$, 0 and $\gamma$ *etc.* and generally

$$w_{abc} = f(a\alpha, b\beta, c\gamma \ldots k\kappa)\alpha\beta\gamma \ldots \kappa \tag{47}$$

molecules with variables p1, p2, ...qr between limits

$$a\alpha \text{ and } (a+1)\alpha, b\beta \text{ and } (b+1)\beta \ldots k\kappa \text{ and } (k+1)\kappa, \tag{48}$$

These limits are so close that we can equate all the values in between, then it is as if the variable $p_1$ could only take the values $0, \alpha, 2\alpha, 3\alpha$ *etc.*, variable $p_2$ could take the values $0, \beta, 2\beta, 3\beta$ *etc.* Let $n$ be the total number of molecules of the first type. We again distinguish the variables for the other gases by the corresponding accents, so that

$$\mathcal{P} = \frac{n!n'!n''!\ldots n^{(v)}!}{\prod w_{abc\ldots k}! \prod w'_{a'b'\ldots k'}! \prod w''_{a''b''\ldots k''}! \ldots \prod w^{(v)}_{a^{(v)}b^{(v)}\ldots k^{(v)}}!} \tag{49}$$

is the possible number of permutations of the elements of this complexion, which we call the permutability. The products are to be read so that the indices $a, b \ldots, a', b' \ldots$ *etc.* run over all possible values, *i.e.*, $-\infty$ to $+\infty$ for orthogonal coordinates, 0 to $2\pi$ for angular coordinates, and so on. Consider first, the case where $p_1$ really can take only the values $0, \alpha, 2\alpha, 3\alpha, \ldots$, and similarly with the other variables; then expression (49) is just the number of complexions this state distribution could have; this number is, according to the assumptions made above a measure of the probability of the state distribution. The variables $w$ and $n$ are all very large; we can again therefore replace $w!$ with $\sqrt{2\pi}(w/e)^w$. We also denote the sum

$$n \ln n + n' \ln n' + \ldots n^{(v)} \ln n^{(v)}$$

by N, so we can also replace $n!$ by $\sqrt{2\pi}(n/e)^n$ and then immediately take the logarithm

$$\ln \mathcal{P} = N - C \ln 2\pi - \left[ \sum w_{ab\ldots} \cdot \ln w_{ab\ldots} + \sum w'_{a'b'\ldots} \cdot \ln w'_{a'b'\ldots} + \ldots \right]. \tag{50}$$

The sums are to be understood in the same sense as the products above. $2C$ is the number of factorials in the denominator of Equation (49) minus $v+1$. Let us now substitute the expression (47) for the variables $w$ into Equation (50) and then take the limit of infinitesimal $\alpha, \beta, \gamma \ldots$.

Omitting unnecessary constants, the magnitude we obtain for the permutability measure, denoted by $\Omega$, is

$$\Omega = -[\int \int \ldots f(p_1, p_2 \ldots q_r) \ln f(p_1, p_2 \ldots q_r)dp_1 dp_2 \ldots dq_r +$$

$$\int \int \ldots f(p'_1, p'_2 \ldots q'_{r'}) \ln f(p'_1, p'_2 \ldots q'_{r'})dp'_1 dp'_2 \ldots dq'_{r'} + \ldots] \tag{51}$$

The integration is to extend over all possible values of the variables. I have in my paper "On the thermal equilibrium of gases on which external forces act." demonstrated that the expression in the square brackets is at a minimum for a gas in a state of thermal equilibrium; including, of course, the kinetic energy constraint equation.

*IV. On the Conditions for the Maximum of the Power-Exponent Free Product Determining the State Distribution Function*

Before I go into the treatment of the second law I want to concisely treat a problem whose importance I believe I have shown in Section I, in the discussion of the work of Mr. Oskar Emil Meyer's on this subject; namely the problem of finding the maximum of the product of the probabilities of all possible states. However I want to deal with this problem only for mono-atomic gases, and with no other constraint than the equation for the kinetic energy. We first consider the simplest case where only a discrete number of kinetic energy values, $0, \epsilon, 2\epsilon \ldots p\epsilon$ are possible, and to start we use kinetic energies, not velocity components, as variables. We again denote by $w_0, w_1, w_2 \ldots w_v$ the number of molecules with kinetic energy $0, \epsilon, 2\epsilon \ldots p\epsilon$.

If we treat the subject in the usual way, the following relationship holds: The quantity

$$B = w_0 \cdot w_1 \cdot w_2 \ldots w_p$$

or, if you prefer, the quantity

$$\ln B = \ln w_0 + \ln w_1 + \ln w_2 + \ldots \ln w_p \tag{52}$$

must be a maximum, with the constraints

$$n = w_0 + w_1 + w_2 + \ldots + w_p \tag{53}$$

and

$$L = (w_1 + 2w_2 + 3w_3 + \ldots + pw_p)\epsilon. \tag{54}$$

If to Equation (52) we add Equation (53) multiplied by $h$, and add Equation (54) multiplied by $k$, then set the partial derivatives of the sum with respect to $w_0, w_1, w_2 \ldots$ equal to zero, we obtain the equations

$$\frac{1}{w_0} + h = 0, \quad \frac{1}{w_1} + h + k = 0, \quad \frac{1}{w_2} + h + 2k = 0 \; etc.$$

from which, by elimination of the constants $h$ and $k$

$$\frac{1}{w_1} - \frac{1}{w_0} = \frac{1}{w_2} - \frac{1}{w_1} = \frac{1}{w_3} - \frac{1}{w_2} = \ldots$$

or

$$\frac{1}{w_0} = a, \quad \frac{1}{w_1} = a + b, \quad \frac{1}{w_2} = a + 2b, \quad \ldots \frac{1}{w_p} = a + pb. \tag{55}$$

Substituting these values into Equations (53) and (54) the two constants $a$ and $b$ can be determined:

$$n = \frac{1}{a} + \frac{1}{a+b} + \frac{1}{a+2b} + \ldots + \frac{1}{a+pb},$$

$$L = \frac{\epsilon}{a+b} + \frac{2\epsilon}{a+2b} + \frac{3\epsilon}{a+3b} + \ldots + \frac{p\epsilon}{a+pb}.$$

The direct determination of the two unknowns $a$ and $b$ from these equations would be extremely lengthy. The method of *Regula falsi* would provide a more rapid solution for each special case; I have not troubled

myself with such calculations, but will give here only a general discussion of how the expected solutions can be easily obtained, keeping in mind that these methods can only provide an approximation solution to the problem, since only positive integers are allowed, but fractional values are not. The first point to note is that the problem ceases to have any meaning as soon as the product $p \cdot (p+1)/2$ is greater than $L/\epsilon$. Because then it necessarily follows that one of the $w's$, and so also the product $B$, is zero.

Then there is no question of a maximum value for B. For the problem to make any sense, an excessive value for the kinetic energy cannot be possible. If

$$p \cdot \frac{p+1}{2} = \frac{L}{\epsilon}$$

then all the $w's$ from $w_0$ onwards must be equal to 1 for $B$ to be non-zero. A greater variation in values can occur only if smaller values of $p$ are chosen. Then, when $n$ is large the above equations provide usable approximations. First, $a$ will be significantly smaller than $b$, so $w_0$ is very large, and $w_1$ will be much smaller; $w_2$ will be close to $w_1/2$, $w_3$ will be close to $2w_2/3$ *etc.* In general, the decrease in the variable $w$ with increasing index will be fairly insignificant when the maximum of $w_0 \cdot w_1 \cdot w_2 \ldots$ is sought, rather than the maximum of $w_0^{w_0} w_1^{w_1} w_2^{w_2} \ldots$. Given much smaller $p$ values, the value of $a$ is not much less than $b$, so $w_0$ is also not that much larger than the other $w's$; then $w_2$ is greater than $w_1/2$, $w_3$ is greater than $(2/3)w_2$ *etc.* The decrease of $w$ with increasing index is even less. Decreasing $p$ still further, $a$ will dominate, and there will be hardly any decrease in $w$ with increasing index. Finally $b$ becomes negative, and the size of $w$ will even increase with increasing index. The following cases provide examples, for each the integer values of the $w's$ which maximize $B$ are given.

$$n = 30, L = 30\epsilon, p = 5, \quad w_0 = 17, w_1 = 5, w_2 = 3, w_3 = 2, w_4 = 2, w_5 = 1.$$
$$n = 31, L = 26\epsilon, p = 4, \quad w_0 = 18, w_1 = 6, w_2 = 3, w_3 = 2, w_4 = 2.$$
$$n = 40, L = 40\epsilon, p = 5, \quad w_0 = 23, w_1 = 7, w_2 = 3, w_3 = 3, w_4 = 2, w_5 = 2.$$
$$n = 40, L = 40\epsilon, p = 6, \quad w_0 = 24, w_1 = 6, w_2 = 3, w_3 = 3, w_4 = 2, w_5 = 1, w_6 = 1.$$
$$n = 18, L = 45\epsilon, p = 5, \quad w_0 = 3, w_1 = 3, w_2 = 3, w_3 = 3, w_4 = 3, w_5 = 3.$$
$$n = 23, L = 86\epsilon, p = 5, \quad w_0 = 1, w_1 = 2, w_2 = 2, w_3 = 3, w_4 = 4, w_5 = 11.$$

Let us now turn to the case where the value of the kinetic energy is continuous; first, consider the kinetic energy $x$ as the independent variable, so the problem, in our view is the following: The expression

$$Q = \int_0^P \ln f(x) dx$$

becomes a maximum, while at the same time

$$n = \int_0^P f(x) dx \quad \text{and} \quad L = \int_0^P x f(x) dx$$

are constant. $P$ is also constant. I have purposefully set the upper integration limit to $P$, not $\infty$. It is then still straightforward to allow $P$ to increase more and more. Proceeding accordingly, we obtain:

$$\delta \int_0^P [\ln x + h f(x) + k x f(x)] dx = \int_0^P \left[ \frac{1}{f} + h + kx \right] dx \delta f = 0$$

from which it follows

$$f = -\frac{1}{h + kx} = \frac{1}{a + bx},$$

if we set $h = -a, k = -b$. To determine these two constants we use the equations

$$n = \int_0^P f(x)dx = \frac{1}{b} \ln \frac{a + bP}{a},$$

$$L = \int_0^P xf(x)dx = \frac{P}{b} - \frac{a}{b_2} \ln \frac{a + bP}{a},$$

which, writing $a/b$ as $\alpha$, leads to:

$$L + \alpha n = \frac{P}{b}, \quad bn = \ln\left(1 + \frac{P}{\alpha}\right)$$

and also

$$(L + \alpha n)\ln\left(1 + \frac{P}{\alpha}\right) = Pn. \tag{56}$$

From this transcendental equation $\alpha$ must be determined, from which it is easy to obtain $a$ and $b$. Since $Pn$ is the kinetic energy which the gas would have if every molecule in it had the maximum possible kinetic energy $P$, we see immediately that $Pn$ is infinitely greater than $L$. $L/n$ is the average kinetic energy of a molecule. It is then easy to verify that $P/\alpha$ cannot be finite because then $Pn/\alpha n$ would be finite, and in the expression $L + \alpha n$ $L$ could be neglected. But then in Equation (56) only $P/\alpha$ terms would remain, and only vanishingly small values of this term could satisfy the equation, which is inconsistent with the original assumption. Nor can $P/\alpha$ be vanishingly small because then $L$ would again be vanishingly small compared to $\alpha n$. Furthermore

$$\ln\left(1 + \frac{P}{\alpha}\right)$$

could be expanded in powers of $P/\alpha$, and the Equation (56) would yield a finite value for $P/\alpha$. There remains only the possibility that $P/\alpha$ is very large. Since

$$\frac{\alpha n}{Pn} \ln \frac{Pn}{\alpha n}$$

vanishes, Equation (56) gives

$$\alpha = \frac{a}{b} = pe^{-\frac{nP}{L}},$$

from which follows:

$$b = \frac{P}{L}, \quad a = \frac{p^2}{L}e^{-\frac{nP}{L}}$$

By the approach used in this section, using the mean kinetic energy of a molecule, these equations show that in the limit of increasing $p, W, L$ the probability of dispersion in kinetic energy remains indeterminate. We now want to consider a second, more realistic problem. We take the three velocity

components $u, v, w$ parallel to the three coordinate axes as the independent variables, and find the maximum of the expression

$$Q = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \ln f(u, v, w) du\, dv\, dw,$$

while simultaneously the two expressions

$$n = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} f(u, v, w) du\, dv\, dw,$$

$$L = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} (u^2 + v^2 + w^2) f(u, v, w) du\, dv\, dw$$

remain constant, integrating over u, v, w. If the velocity magnitude is

$$\omega = \sqrt{u^2 + v^2 + w^2}$$

and its direction is given by the two angles $\theta$ and $\varphi$ (length and breadth [13]), we have as is well known

$$du\, dv\, dw = \omega d\omega \sin \theta d\theta d\varphi,$$

hence

$$Q = 4\pi \int\limits_0^P \ln f(\varrho) \varrho^2 d\varrho, \tag{57}$$

$$n = 4\pi \int\limits_0^P \varrho^2 f(\varrho) d\varrho, \tag{58}$$

$$L = 4\pi \int\limits_0^P \varrho^4 f(\varrho) d\varrho \tag{59}$$

If there are no external forces, then clearly $f(u, v, w)$ is independent of direction of the velocity. Instead of integrating to infinity, we intentionally integrate to a finite value of $P$. Evaluating $f(\varrho)$ just as we did for $f(x)$ earlier, we obtain

$$f(\varrho) = -\frac{1}{h + k\varrho^2} = \frac{1}{a^2 + b^2 \varrho^2},$$

where we set $-h = a^2$ and $-k = b^2$ . The two constant $a$ and $b$ are to be determined from Equations (58) and (59) which become, given the value of $f(\varrho)$:

$$n = 4\pi \int\limits_0^P \frac{\varrho^2 d\varrho}{a^2 + b^2 \varrho^2} = 4\pi \left( \frac{P}{b^2} - \frac{a}{b^3} arctg \frac{bP}{a} \right)$$

---

[13]   Altitude and azimuth?

$$L = 4\pi \int\limits_0^P \frac{\varrho^4 d\varrho}{a^2 + b^2 \varrho^2} = \frac{4\pi P^3}{3b^2} - \frac{a^2}{b^2} n. \tag{60}$$

From the last equation we get

$$\frac{4\pi}{b^2} = \frac{3}{P^3} \left( L + \frac{a^2}{b^2} n \right).$$

Substituting this equation into the first equation of (60) gives

$$n = 3 \left( \frac{L}{P^2} + \frac{a^2}{b^2 P^2} n \right) \left( 1 - \frac{a}{Pb} arctg \frac{bP}{a} \right). \tag{60a}$$

If however $b_2$ is negative, we put $-b_2$ instead of $b_2$ and obtain:

$$f(\varrho) = \frac{1}{a^2 - b^2 \varrho^2},$$

$$n = 4\pi \left( -\frac{P}{b^2} + \frac{a}{2b^3} \ln \left( \frac{a + bP}{a - bP} \right) \right)$$

$$L = -\frac{4\pi P^3}{3b^2} + \frac{a^2}{b^2} n.$$

$$n = 3 \left( \frac{L}{P^2} - \frac{a^2 n}{b^2 P^2} \right) \left( 1 - \frac{a}{2bP} \ln \left( \frac{a + bP}{a - bP} \right) \right). \tag{60b}$$

From Equations (60a) and (60b) one first has to calculate the ratio $a/b$; and by the same means by which Equation (56) was analyzed, we first determine whether $bP/a$ is infinitely small, finite, or infinitely large, the only difference being in Equation (60a) every infinitesimal variation of $L/nP^2$ occurs. However I will not discuss the point further, except to note that as $n$ and $P$ grow larger, one also cannot get a result which depends only on the average kinetic energy. Also I will not discuss in detail those cases where there are other constraint equations besides the equation for kinetic energy, as this would lead me too far afield.

To provide a demonstration of how general the concept of the most probable state distribution of gas molecules is, here I supply another definition for it. Suppose again that each molecule can only have a discrete number of values for the kinetic energy, $0, \epsilon, 2\epsilon, 3\epsilon \ldots \infty$.

The total kinetic energy is $L = \lambda\epsilon$. We want to determine the kinetic energy of each molecule in the following manner: We have in an urn just as many identical balls $(n)$ as molecules present. Every ball corresponds to a certain molecule. We now make $\lambda$ draws from this urn, returning the ball to the urn each time. The kinetic energy of the first molecule is now equal to the product of $\epsilon$ and the number of times the ball corresponding to this molecule is drawn. The kinetic energies of all other molecules are determined analogously. We have produced a distribution of the kinetic energy $L$ among the molecules (a complexion). We again make $\lambda$ draws from the urn and produce a second complexion, then a third, *etc.* many times $(J)$, and produce $J$ complexions. We can define the most probable state distribution in two ways: First, we find how often in all $J$ complexions a molecule has kinetic energy 0, how often the kinetic energy is $\epsilon, 2\epsilon$, *etc.*, and say that the ratios of these numbers should provide the probabilities that a molecule has kinetic energy $0, \epsilon, 2\epsilon$, *etc.* at thermal equilibrium. Second, for each complexion we form the corresponding state distribution. If some state distribution is composed of $\mathcal{P}$ complexions, we

then denote the quotient $\mathcal{P}/J$ as the probability of the state distribution. At first glance this definition of a state distribution seems very plausible. But we shall presently see that this should not be used, because under these conditions, the distribution whose probability is the greatest would not correspond to thermal equilibrium. It is easy to cast the hypothesis that concerns us into formulas. First of all, we want to discuss the first method of probability determination. We consider the first molecule, and assume that $\lambda$ draws were made; the probability that the first molecule was picked in the first draw is $1/n$; however the probability that another ball was drawn is $(n-1)/n$. Thus the probability that on the $1^{st}, 2^{nd}, 3^{rd} \ldots k^{th}$ draws the molecule corresponding to the first ball has been picked, and then a different ball for each of the following, is given by

$$\left(\frac{1}{n}\right)^k \left(\frac{n-1}{n}\right)^{\lambda-k} = \left(\frac{n-1}{n}\right)^\lambda \cdot \left(\frac{1}{n-1}\right)^k$$

Likewise is the probability that the ball corresponding to the first molecule is picked on the $1^{st}, 2^{nd}, 3^{rd} \ldots (k-1)^{th}$, and then (k+1)th draws *etc.* The probability that the ball corresponding to the first molecule is picked for any arbitrary $k$ draws, and not for the others is

$$w_k = \frac{\lambda!}{(\lambda-k)!k!} \left(\frac{n-1}{n}\right)^\lambda \left(\frac{1}{n-1}\right)^k .$$

This probability, that a molecule has then kinetic energy $k\epsilon$ is exactly the same for all the other molecules. Using again the approximation formula for the factorial, we obtain

$$w_k = \sqrt{\frac{1}{2\pi}} \cdot \left(\lambda\frac{n-1}{n}\right)^\lambda \cdot \frac{1}{\sqrt{\left(1-\frac{k}{\lambda}\right)\cdot k}} \cdot \left[\frac{\lambda-k}{(n-1)k}\right]^k \cdot (\lambda-k)^{-\lambda},$$

which shows that the probability of the larger kinetic energies is so disproportionately important that the entire expression does not approach a clearly identifiable limit with increasing $k, \lambda, 1/\epsilon$ and $n$. We will now proceed to the second possible definition of the most probable state distribution. We need to consider all $J$ complexions that we have formed by $J$ drawings of $\lambda$ balls from our urn. One of the various possible complexions consists of $\lambda$ drawings of the ball corresponding to the first ball. We want to express this complexion symbolically by $m_1^\lambda \cdot m_2^0 \cdot m_3^0 \ldots m_n^0$. A second complexion, with $\lambda-1$ draws of the ball corresponding to the first molecule, and one draw of the ball corresponding to the second molecule we want to express as

$$m_1^{\lambda-1} \cdot m_2^1 \cdot m_3^0 \ldots m_n^0$$

We see that the different possible complexions are expressed exactly by various components; the sum of these appears as the power series

$$(m_1 + m_2 + m_3 + \ldots + m_n)^\lambda \tag{A}$$

that is developed according to the polynomial theorem. The probability of each such complexion is thus exactly proportional to the coefficient of the corresponding power series term, when you first form the product

$$(m_1' + m_2' + \ldots + m_n')(m_1'' + m_2'' + \ldots + m_n'') \ldots (m_1^{(\lambda)} + m_2^{(\lambda)} + + \ldots + m_n^{(\lambda)})$$

and finally omit from this product the upper indexes, which then generates a term exactly proportional to the polynomial coefficient. Then by the symbol $m_1' \cdot m_3'' \cdot m_7''' \ldots$ we understand that the first pick corresponded to the first molecule, the second pick corresponded to the third molecule, on the third pick the ball corresponding to the seventh molecule was picked out, *etc.* All possible products of the variables $m_1', m_1'', m_2'$ *etc.* represent equi-probable complexions. We want to know how often among all the terms of the power series (A) (whose total number is $n^\lambda$ ), there occur terms whose coefficients contain any one state distribution. For example, consider the state distribution where one molecule has all the kinetic energy, all others have zero kinetic energy. This state distribution appears to correspond to the following members of the power series (A)

$$m_1^\lambda \cdot m_2^0 \cdot m_3^0 \ldots, m_1^0 \cdot m_2^\lambda \cdot m_3^0 \ldots, m_1^0 \cdot m_2^0 \cdot m_3^\lambda \ldots \quad etc.$$

with 'undivided' $\lambda$. Similarly for the state distribution in which $w_0$ molecules have kinetic energy zero, $w_1$ molecules have kinetic energy $\epsilon$, $w_2$ molecules have kinetic energy $2\epsilon$, *etc.* there are

$$\frac{\lambda!}{w_0! w_1! w_2! \ldots w_\lambda!}$$

members of the power series (A). Each of these elements has the same polynomial coefficient, and that is identical to

$$\frac{\lambda!}{(0!)^{w_0}.(1!)^{w_1}.(2!)^{w_2} \ldots (\lambda!)^{w_\lambda}}$$

In summary, therefore, according to the now accepted definition, the probability of this state distribution is

$$\frac{(\lambda!)^2}{n^\lambda} \cdot \frac{1}{w_0! w_1! w_2! \ldots w_\lambda!} \cdot \frac{1}{(0!)^{w_0}.(1!)^{w_1}.(2!)^{w_2} \ldots (\lambda!)^{w_\lambda}}$$

However, the maximization of this quantity also does not lead to the state distribution corresponding to thermal equilibrium.

### *V. Relationship of the Entropy to That Quantity Which I Have Called the Probability Distribution*

When considering this relationship, let us initially deal with the simplest and clearest case, by first investigating a monoatomic gas on which no external forces act. In this case formula (34) of Section II applies. To give it full generality, however, this formula must also include the $x, y, z$ coordinates of the position of the molecule. The maximum of such a generalized expression (34) then yields, not only the distribution of the velocity components of the gas molecules, which was sufficient for the case considered there, but also the distribution of the whole mass of gas in an enclosing vessel, where there it was taken for granted that the gas mass fills the vessel uniformly.

The generalization of (34) for the permutability measure can be easily obtained from Equation (51) by substituting $x, y, z, u, v, w$ for $p_1, p_2 \ldots q_r$ and simply omitting the terms with accented variables. It reads as follows:

$$\Omega = - \int \int \int \int \int \int f(x, y, z, u, v, w) \ln f(x, y, z, u, v, w) dx dy dz du dv dw, \qquad (61)$$

where $f(x, y, z, u, v, w) dx dy dz du dv dw$ is the number of gas molecules present for which the six variables $x, y, z, u, v, w$ lie between the limits $x$ and $x + dx$, $y$ and $y + dy$, $z$ and $z + dz \ldots$ *etc.*, w

and $w + dw$ and the integration limits for the velocity are between $-\infty$ and $+\infty$, and for the coordinates over the dimensions of the vessel in which the gas exists. If the gas was not previously in thermal equilibrium, this quantity must grow. We want to compute the value this quantity has when the gas has reached the state of thermal equilibrium. Let $V$ be the total volume of the gas, $T$ be the average kinetic energy of a gas molecule, and $N$ be the total number of molecules of the gas; finally $m$ is the mass of a gas molecule. There is then for the state of thermal equilibrium

$$f(x, y, z, u, v, w) = \frac{N}{V \left(\frac{4\pi T}{3m}\right)^{\frac{3}{2}}} \cdot e^{-\frac{3m}{4T}(u^2 + v^2 + w^2)}.$$

Substituting this value into Equation (61) gives

$$\Omega = \frac{3N}{2} + N \ln \left[ V \left(\frac{4\pi T}{3m}\right)^{\frac{3}{2}} \right] - N \ln N \tag{62}$$

If by $dQ$ we denote the differential heat supplied to the gas, where

$$dQ = NdT + pdV \tag{63}$$

and

$$pV = \frac{2N}{3} \cdot T. \tag{64}$$

p is the pressure per unit area, the entropy of the gas is then:

$$\int \frac{dQ}{T} = \frac{2}{3} N \cdot \ln \left( V \cdot T^{\frac{3}{2}} \right) + C.$$

since N is regarded as a constant, with a suitable choice of constant

$$\int \frac{dQ}{T} = \frac{2}{3}\Omega. \tag{65}$$

It follows that for each so-called reversible change of state, wherein in the infinitesimal limit the gas stays in equilibrium throughout the change of state, the increase in the Permutability measure $\Omega$ multiplied by $\frac{2}{3}$ is equal to $\int dQ/T$ taken over the change in state, *i.e.*, it is equal to the entropy increment. Whereas in fact when a very small amount of heat $dQ$ is supplied to a gas, its temperature and volume increase by $dT$ and $dV$. Then it follows from Equations (63) and (64)

$$dQ = NdT + \frac{2N}{3V} \cdot TdV$$

while from Equation (62) it is found that:

$$d\Omega = +N\frac{dV}{V} + \frac{3N}{2} \cdot \frac{dT}{T}.$$

It is known that if in a system of bodies many reversible changes are taking place, then the total sum of the entropy of all these bodies remains constant. On the other hand, if among these processes are ones that are not reversible, then the total entropy of all the bodies must necessarily increase, as is well known from the fact that $\int dQ/T$ integrated over a non-reversible cyclic process is negative.

According to Equation (65) the sum of the permutability measures of the bodies $\Sigma\Omega$ and the total permutability measure must have the same increase. Therefore at thermal equilibrium the magnitude of the permutability measure times a constant is identical to the entropy, to within an additive constant; but it also retains meaning during a non-reversible process, continually increasing.

We can establish two principles: The first refers to a collection of bodies in which various state changes have occurred, at least some of which are irreversible; e.g., where some of the bodies were not always in thermal equilibrium. If the system was in a state of thermal equilibrium before and after all these changes, then the sum of the entropies of all the bodies can be calculated before and after those state changes without further ado, and it is always equal to 2/3 times the permutability measure of all the bodies. The first principle is that the total entropy after the state changes is always greater than before it. The same is of course true of the permutability measure. The second principle refers to a gas that undergoes a change of state without requiring that it begin and end in thermal equilibrium. Then the entropy of the initial and final states is not defined, but one can still calculate the quantity which we have called the permutability measure; and that is to say that its value is necessarily larger after the state change than before. We shall presently see that the latter proposition can be applied without difficulty to a system of several gases, and it can be extended as well to polyatomic gas molecules, and when external forces are acting. For a system of several gases, the sum of the individual gas permutability measures must be defined to be the permutability measure of the system; if one introduces on the other hand the number of permutations itself, then the number of permutations of the system would be the product of the number of permutations of the constituents. If we assume the latter principle applies to any body, then the two propositions just discussed are special cases of a single general theorem, which reads as follows:

*We consider any system of bodies that undergoes some state changes, without requiring the initial and final states to be in thermal equilibrium. Then the total permutability measure for the bodies continually increases during the state changes, and can remain constant only so long as all the bodies during the change of state remain infinitely close to thermal equilibrium (reversible state changes.)*

To give an example, we consider a vessel divided into two halves by a very thin partition. The remaining walls of the vessel should also be very thin, so that the heat they absorb can be neglected, and surrounded by a substantial mass of other gas. One half of the vessel should be completely filled with gas, while the other is initially completely empty. Suddenly pulling away the divider, which requires no significant work, causes that gas to spread at once throughout the vessel. Calculating the permutability measure for the gas, we find that this increases during this process, without changes in any other body. Now the gas is compressed in a reversible manner to its old volume by a piston. To achieve this manipulation , we can if we want assume that the piston is created by a surrounding dense gas enclosed in infinitely thin walls. This gas is unchanged except that it moves down in space.

Since the permutability measure does not depend on the absolute position in space, the permutability measure of the gas driving the piston does not change. That of the gas inside the vessel decreases to the initial value, since this gas has gone through a cyclic process. However, since this was not reversible, $\int dQ/T$ integrated over this cycle is not equal to the difference between the initial and final values of the entropy, but it is smaller due to the uncompensated transformation in the expansion. In contrast, heat is transferred to the surrounding gas. So for this process the permutability measure of the surrounding gas

is increased by as much as that of the enclosed gas in the vessel during the first process. Since the latter mass of gas went through a cyclic process, its entropy decreased by as much during the second process as it increased during the first, but not by $\int dQ/T$; and because the second process was reversible, the entropy of the surrounding gas increased as much as the enclosed gas's decreased. The result is, as it has to be, that the sum of the permutability measures of all bodies of gas has increased. For a gas which moves at a constant speed in the direction of the x-axis [14]

$$f(x, y, z, u, v, w) = V \frac{N}{\sqrt{\left(\frac{4\pi T}{3m}\right)^3}} \cdot e^{-\frac{3m}{4T}((u-\alpha)^2 + v^2 + w^2)}. \tag{66}$$

If we substitute this expression into Equation (61) we get exactly Equation (62) again. Thus the translational movement of a mass of gas does not increase its permutability measure. And the same is true for the kinetic energy arising from any other net mass movement (molar movement); because it arises from the progression of the individual volume elements and their deformations and rotations which are of a higher order, - infinitely small - and therefore entirely negligible. Here we obviously ignore the changes of permutability measure due to internal friction or temperature changes connected with those molecular motions. The temperature $T$ of the moving gas is understood to mean half of the average value of $m[(u - \alpha)^2 + v^2 + w^2]$. So if frictional and temperature changes are not present, (e.g., if a gas, together with its enclosing vessel falls freely) a net mass movement has no effect on the permutability measure, until its kinetic energy is converted into heat; which is why molar motion is known as heat of infinite temperature.

Let's now move on to a mono-atomic gas on which gravity acts. The permutability measure is represented by the same equation (51), but instead of the generalized coordinates we again introduce $x, y, z, u, v, w$. Equation (51) thus gives us a value for $\Omega$ which is exactly the same as Equation (61). In the case of thermal equilibrium one has

$$f(x, y, z, u, v, w) = Ce^{-\frac{3}{2T}\left(gz + \frac{m\omega^2}{2}\right)}, \tag{67}$$

where $\omega^2 = u^2 + v^2 + w^2$. The constant C is determined by the density of the gas. One has e.g., a prismatic shaped vessel of height $h$, with a flat, horizontal bottom surface with area $= q$. Further, let $N$ be the total number of gas molecules in the vessel, and $z$ denote the height of a gas molecule from the bottom of the vessel, then

$$C = \frac{N}{\left(\frac{4\pi T}{3m}\right)^{\frac{3}{2}} \cdot q \int_0^h e^{-\frac{3gz}{2T}} dz} = \frac{N}{\left(\frac{4\pi T}{3m}\right)^{\frac{3}{2}} \cdot q \cdot \frac{2T}{3g}\left(1 - e^{-\frac{3gh}{2T}}\right)}, \tag{68}$$

from which it follows:

$$\Omega = \frac{3N}{2} + N\ln\left(\frac{4\pi T}{3m}\right)^{\frac{3}{2}} + N\ln q + N\ln\frac{2T}{3g}\left(1 - e^{-\frac{3gh}{2T}}\right) + N\left(1 - \frac{3ghe^{-\frac{3gh}{2T}}}{2T\left(1 - e^{\frac{3gh}{2T}}\right)}\right) - N\ln N. \tag{69}$$

---

[14]  Here $V$ should appear in the denominator, cf the equation after (61).

It is immediately apparent from the last formula that when a mass of gas falls a bit lower, without otherwise undergoing a change, that does not change its $\Omega$ value a bit. (Of course gravity acts as a constant downward force, and its increase with the approach to the earth's center is neglected, which is always the case in heat theory problems.)

Let's go now to the most general case of an arbitrary gas on which any external force acts, so we again apply Equation (51). However, in order that the formulas not be too lengthy, let only one type of gas be present in the vessel. The permutability measure of a gas mixture can then be found without difficulty, since it is simply equal to the sum of the permutability measures each component would have if present in the vessel alone. For thermal equilibrium, then

$$f = \frac{Ne^{-h(\chi+L)}}{\int\int e^{-h(\chi+L)}dodw},$$

where $\chi$ is the potential energy, $L$ is the kinetic energy of a molecule, $N$ is the number of molecules in the vessel, where

$$do = dp_1 dp_2 \ldots dp_r, \quad dw = dq_1 dq_2 \ldots dq_r$$

It therefore follows:

$$\Omega = -\int\int f\ln f\, dodw = -N\ln N + N\ln\int\int e^{-h(\chi+L)}dodw + hN\bar\chi + \frac{rN}{2}. \tag{70}$$

In the penultimate term, $\bar\chi$ is the average potential energy of a molecule, whose magnitude is

$$\frac{1}{N}\int\int \chi f\, dodw = \frac{\int\int \chi e^{-h(\chi+L)}dodw}{\int\int e^{-h(\chi+L)}dodw}$$

The last term can be found by taking into account that

$$L = \frac{\int\int Le^{-h(\chi+L)}dodw}{\int\int e^{-h(\chi+L)}dodw} = \frac{r}{2h}$$

(In respect of the above, see the already cited book of Watson, pp 36 and 37). The second term on the right of Equation (70) can be further transformed, if one introduces instead of $q_1, q_2 \ldots q_r$ the variables $s_1, s_2 \ldots s_r$, which have the property that the term $L$ is reduced to $s_1^2 + s_2^2 + \ldots + s_r^2$. We then denote by $\Delta$ the Jacobian

$$\sum \pm \frac{dq_1}{ds_1}\frac{dq_2}{ds_2}\cdots\frac{dq_r}{ds_r},$$

So then

$$\int\int e^{-h(\chi+L)}dodw = \left(\frac{\pi}{h}\right)^{\frac{r}{2}}\int \Delta e^{-h\chi}do, \quad \bar\chi = \frac{\int \Delta\chi e^{-h\chi}do}{\int \Delta e^{-h\chi}do}$$

and so also

$$\Omega = N\ln\int \Delta e^{-h\chi}do - \frac{Nr}{2}\ln h + hN\bar\chi + \frac{rN}{2}(1+\ln\pi) - N\ln N. \tag{71}$$

To this expression, one can compare Equation (18) of my paper "Analytical proof of the second law of thermodynamics from the approach to equilibrium in kinetic energy" (Wiss. Abhand. Vol I, reprint 20), or Equation (95) of my "Further Studies" (Wiss. Abhand. Vol I, reprint 22) by replacing $p_1, p_2 \ldots$ with $x_1, y_2 \ldots$; $q_1, q_2 \ldots$ with $u_1, v_2 \ldots$; $s_1, s_2 \ldots$ with

$$\sqrt{\frac{m}{2}}u_1, \quad \sqrt{\frac{m}{2}}v_1 \ldots,$$

$r$ with $3r$, whereby

$$\Delta = \left(\frac{2}{m}\right)^{\frac{3r}{2}}$$

It can be seen that Equation (71) is identical to $3N/2$ times Equation (18) of the aforementioned paper, except for an additive constant, wherein multiplying by $N$ indicates that Equation (18) applies to one molecule only. Equation (95) of "Further Studies" is in an opposite manner denoted as $\Omega$ and thus also as the entropy. Taken with a negative sign, it is however greater than $\Omega$ here by $N \ln N$. The former is because in the "Further Studies" I was looking for a value which must decrease, as a result of this, introducing the magnitude of $f^*$ instead of using $f$; this was however less clear. From this agreement it follows that our statement about the relationship of entropy to the permutability measure applies to the general case exactly as it does to a monatomic gas.

Up to this point, these propositions may be demonstrated exactly using the theory of gases. If one tries, however, to generalize to liquid drops and solid bodies, one must dispense with an exact treatment from the outset, since far too little is known about the nature of the latter states of matter, and the mathematical theory is barely developed. But I have already mentioned reasons in previous papers, in virtue of which it is likely that for these two aggregate states, the thermal equilibrium is achieved when Equation (51) becomes a maximum, and that when thermal equilibrium exists, the entropy is given by the same expression. It can therefore be described as likely that the validity of the principle which I have developed is not just limited to gases, but that the same constitutes a general natural law applicable to solid bodies and liquid droplets, although the exact mathematical treatment of these cases still seems to encounter extraordinary difficulties.

## Acknowledgements

## Author Contributions

Both authors contributed equally to the translation.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Braun, S.; Ronzheimer, J.P.; Schreiber, M.; Hodgman, S.S.; Rom, T.; Bloch, I.; Schneider, U. Negative Absolute Temperature for Motional Degrees of Freedom. *Science* **2013**, *339*, 52–55.
2. Dunkel, J.; Hilbert, S. Consistent thermostatistics forbids negative absolute temperatures. *Nat. Phys.* **2014**, *10*, 67–72.
3. Frenkel, D.; Warren, P.B. Gibbs, Boltzmann, and negative temperatures. **2014**, arXiv:1403.4299v1.
4. Vilar, J.M.G.; Rubi, J.M. Communication: System-size scaling of Boltzmann and alternate Gibbs entropies. *J. Chem. Phys.* **2014**, *140*, 20101.

5. Boltzmann, L. Ableitung des Stefanschen Gesetzes, betreffend die Abhängigkeit der Wärmestrahlung von der Temperatur aus der elektromagnetischen Lichttheorie. *Wiedemann's Annalen* **1884**, *22*, 291–294.

6. Planck, M. Section 134: Entropie und Warscheinlichkeit. In *Vorlesungen uber die theorie der wurmestrahlung*; J.A. Barth: Leipzig, Germany, 1906.

7. Van Kampen, N.G. The Gibbs Paradox. In *Essays in Theoretical Physics in Honour of Dirk ter Haar*; Parry, W.E., Ed.; Pergamon: Oxford, UK, 1984; pp. 303–312.

8. Münster, A. 1912. *Statistical Thermodynamics*; Springer: Berlin, Germany, 1969.

9. Ehrenfest, P.; Trkal, V. Deduction of the dissociation-equilibrium from the theory of quanta and a calculation of the chemical constant based on this. *Ann. Phys.* **1921**, *65*, 609–628.

10. Jaynes, E. The Gibbs Paradox. In *Maximum Entropy and Bayesian Methods*; Smith, C., Erickson, G., Neudorfer, P., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1992; pp. 1–22.

11. Swendsen, R.H. Statistical mechanics of colloids and Boltzmann's definition of the entropy. *Am. J. Phys.* **2006**, *74*, 187–190.

12. Pauli, W. *Statistical Mechanics*; MIT Press: Cambridge, MA, USA, 1973.

13. Gallavotti, G. *Statistical Mechanics: A Short Treatise*; Springer: Berlin, Germany, 1999.

14. Gibbs, J.W. *Elementary Principles in Statistical Mechanics*; Dover Publications: New York, NY, USA and London, UK, 1902.

15. Lebowitz, J. Boltzmann's entropy and time's arrow. *Phys. Today* **1993**, *46*, pp. 32–38.

16. Cercignani, C. *Ludwig Boltzmann: The Man Who Trusted Atoms*; Oxford Univ. Press: Oxford, UK, 1998.