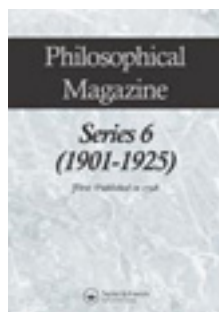


This article was downloaded by: [Cornell University]

On: 03 July 2012, At: 09:48

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number:
1072954 Registered office: Mortimer House, 37-41 Mortimer Street,
London W1T 3JH, UK



Philosophical Magazine Series 6

Publication details, including instructions
for authors and subscription information:
<http://www.tandfonline.com/loi/tphm17>

XLIV. On the partition of energy

C.G. Darwin M.A.F.R.S. ^a & R.H. Fowler M.A.
^b

^a Christ's College, Camb.

^b Trinity College, Camb.

Version of record first published: 08 Apr
2009

To cite this article: C.G. Darwin M.A.F.R.S. & R.H. Fowler M.A. (1922): XLIV.
On the partition of energy , Philosophical Magazine Series 6, 44:261, 450-479

To link to this article: <http://dx.doi.org/10.1080/14786440908565189>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug

doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

XLIV. *On the Partition of Energy.* By C. G. DARWIN, M.A., F.R.S., Fellow and Lecturer in Christ's College, Camb., and R. H. FOWLER, M.A., Fellow and Lecturer in Trinity College, Camb.*

§ 1. *Introduction.*

AN important branch of atomic theory is the study of the way in which energy is partitioned among an assembly of a large number of systems—molecules, Planck vibrators, etc. This study is based on the use of the principles of probability which show that one type of arrangement is much more common than any other. The most usual method is to obtain an expression for the probability of any state described statistically and then to make this probability a maximum. This always involves a use of Stirling's approximation for factorials, which in many cases is illegitimate at first sight, and though it is possible to justify it subsequently, this justification is quite troublesome. It is also usually required to find the relation of the partition to the temperature rather than to the total energy of the assembly, and this is done by means of Boltzmann's theorem relating entropy to probability—a process entailing the same unjustified approximations.

The object of the present paper is to show that these calculations can all be much simplified by examining the average state of the assembly instead of its most probable state. The two are actually the same, but whereas the most probable state is only found by the use of Stirling's formula, the average state can be found rigorously by the help of the multinomial theorem, together with a certain not very difficult theorem in the theory of the complex variable. By this process it is possible to evaluate the average energy of any group in the assembly, and hence to deduce the relation of the partition to temperature, without the intermediary of entropy. The temperature here is measured on a special scale, which can be most simply related to the absolute scale by the use of the theorem of equipartition, and we shall also establish the same relationship directly by connecting it with the scale of a gas thermometer. Throughout the paper the analysis is presented with some attempt at rigour, but it will be found that apart from this rigour it is exceedingly easy to apply the method of calculation. Most of the results are not

* Communicated by the Authors.

new; it is the point of view and the method which, we think, differ from previous treatments

No discussion of the question of partition would be complete without consideration of its relation to thermodynamic principles. We shall leave this view of the subject to a future paper; for the increased light thrown on the statistical nature of entropy raises many interesting points which could not be discussed here properly without making the present work run to inordinate length.

§ 2. *Statistical Principles and Weight.*

Before proceeding to the problem it will be well to review, in general outline, the principles of the theory of the partition of energy, though we have nothing new to say in this connexion. We shall be concerned with collections of molecules, Planck vibrators, etc.—each individual unit will be called a *system*, and we shall call the whole collection an *assembly*. We shall be dealing mainly with assemblies composed of groups of systems, the individuals in each group being identical in nature. In order to make the problem definite it is necessary to assume that each system has a definite assignable energy, and yet can interact with the others. This requires that the time of interactions, during which there will be energy which cannot be assigned to a definite single system, is negligibly small compared with the time during which each system describes its own motion.

For such an assembly we are to calculate various average properties of its state, when it describes its natural motion according to whatever laws it may obey. There will at any rate be an energy integral, and we have therefore to calculate these averages subject to the condition of constant energy. To determine the basis on which these averages are to be calculated we are to apply the principles of probability; and the calculation of itself falls into two stages, the prior and the statistical. The prior stage aims at establishing what are the states which are to be taken as of equal probability. In the statistical stage we have simply to enumerate the states specified in the prior stage, allow for the fact that the systems are macroscopically indistinguishable, and evaluate the averages taken over these states.

It is not here our purpose to enter into a full discussion of the fundamental questions that arise in connexion with the determination of what states ought to be taken as equally

probable. It will suffice to recall that for assemblies obeying the laws of classical mechanics the theorem of Liouville shows that the elements of equal probability may be taken to be equal elements of volume in Gibbs' "phase space." It follows out of this, for example, in the case of an assembly of a number of identical systems—say simple free molecules—that the elements of equal probability can be simplified down into 6-dimensional cells $dq_1 dq_2 dq_3 dp_1 dp_2 dp_3$ of equal extension, where q_1, q_2, q_3 are the coordinates, and p_1, p_2, p_3 the conjugated momenta, of a single molecule. We shall describe this by saying that the *weight* of every equal element $dq_1 \dots dp_3$ is the same, and by a slight generalization, that the weights of unequal cells are proportional to their 6-dimensional extension. The word *weight* is here used in exactly the sense of the term *a priori probability*, as used by Bohr and others.

But when we come to the quantum theory, mechanical principles cease to hold, and we require a new basis for assigning the equally probable elements. Such a basis is provided by Ehrenfest's* Adiabatic Hypothesis and Bohr's† Correspondence Principle. These show how the theorem of Liouville is to be extended, and allow us to assign a *weight* for each quantized state of a system. It is found that we must assign an equal weight to every permissible state in each quantized degree of freedom. At first sight this is a little surprising, for it would seem natural to suppose that a vibrator which could only take energy in large units would be less likely to have a unit than one which could take it in small; but this is to confuse the two stages of the problem. It is only by the supposition of equal weights that we can obtain consistency with classical mechanics by the Correspondence Principle. It is customary‡ in assigning a definite weight to every quantized state to give it the value h , so as to bring the result to the same dimensions as those of the element $dq dp$ in the classical case. But there is considerable advantage in reversing this, and taking the quantized weights as unity and the weight of the element in the phase space as $dq dp/h$; for if this is done, the arguments about entropy are simplified by the absence of logarithms of dimensional quantities. We shall adopt this convention here, though in

* Ehrenfest, Proc. Acad. Amst. xvi. p. 591; Phil. Mag. xxxiii. p. 500 (1917), etc.

† Bohr, "The Quantum Theory of Line Spectra," Dan. Acad. iv. p. 1 (1918).

‡ Ehrenfest & Trkal, Proc. Acad. Sc. xxiii. p. 162. See in particular p. 165 and Additional Notes, No. 1.

all our results it is immaterial—indeed, until such questions as dissociation are considered it makes no difference to adopt different conventions for different types of system. The convention has the advantage of shortening a good many formulæ and freeing them from factors which are without effect on the final results.

An exception to the above rule for assigning weights to quantized motions occurs in the case of *degenerate systems*, where there are two degrees of freedom possessing the same or commensurable frequencies. In this case there is only one quantum number, and the state of the system is partly arbitrary. Bohr* shows that the rational generalization is to assign to such a state a weight factor which can be evaluated by treating the system as the limit of a non-degenerate system, and quantizing it according to any pair of variables in which it is possible to do so. The number of the permissible states which possess the same total quantum number will give the weight of the state. A corresponding rule holds for systems degenerate in three or more degrees of freedom.

The meaning of weight can perhaps be made clearer by considering its introduction the other way round—beginning with an assembly of simple quantized systems of various frequencies. Given the energy, there is a definite number of possible states, which are fully specified by the energy assigned to each system. We then make the hypothesis that it is right to assign an equal probability to each such state in the calculation of averages. This is now the fundamental postulate. The generalization to degenerate systems goes as before, by introducing weight factors. Finally, passing over to mechanical systems, such as free molecules, we are led by an appeal to the converse of the Correspondence Principle to attach weight $dq_1 \dots dp_3/h^3$ to each 6-dimensional cell which specifies completely the state of a single molecule.

The second, statistical, half of the problem consists in enumerating the various *complexions* possible to the assembly. By a *complexion* we mean every arrangement of the assembly, in which we are supposed to be able to distinguish the individuality of the separate systems. We count up the total number of complexions which conform to any specified statistical state of the assembly, and attach to each the appropriate weight factor. Thus the probability of this state is the ratio of the number of its weighted complexions to the

* Bohr, *loc. cit.* p. 26.

total number of all possible weighted complexions. This part of the problem depends on the nature of the particular assembly considered, and so must be treated separately in each case.

We start in § 3 with a problem which concerns not the partition of energy, but the distribution of molecules in a volume. It illustrates the method in its simplest aspect and has the advantage of being purely algebraic. Next, in § 4, we take the distribution of energy among a set of similar Planck vibrators, which is again a purely algebraic process, and then proceed in § 5 to introduce the main theme of this paper by dealing with the partition of energy between two sets of Planck vibrators of different type. This is most conveniently treated by using the complex variable, and in § 6 there is a discussion of the required theorem. The partition of itself introduces the temperature, and in § 7 the special scale is compared with the absolute. In §§ 9, 10, 11 the partition law is generalized to more complicated types of system, such as the quantized rotations of molecules. In §§ 12, 13 the method is extended so as to deal with the free motion of monatomic molecules, intermixed with vibrators. The work leads to a rather neat method of establishing the Maxwell distribution law.

§ 3. *The Distribution of Molecules in Space.*

The first example we shall take is not one of a partition of energy, but of the distribution of small molecules in a vessel. It illustrates in its simplest form the averaging process, and has the advantage of depending only on elementary algebra.

Let there be M molecules, and divide the vessel into m cells of equal or unequal volumes $v_1, v_2 \dots v_m$, which may each be as large or as small as we like. Then

$$v_1 + v_2 + \dots + v_m = V. \quad (3.1)$$

By well-known arguments which we need not consider, it follows that any one molecule is as likely to be in any element of volume as in any other equal one. So by a slight extension of the idea of weight we attach weights $v_1, v_2, \dots v_m$ to the cells. To specify the statistical state we say that the first cell has a_1 molecules, the second a_2 , and so on. Then

$$a_1 + a_2 + \dots + a_m = M. \quad (3.2)$$

By the theory of permutations the number of complexions

which conform to the specification is

$$\frac{M!}{a_1! a_2! \dots a_m!},$$

and each of these must be weighted with a factor

$$v_1^{a_1} v_2^{a_2} \dots v_m^{a_m}.$$

The total number of all the weighted complexions is

$$C = \sum_a \frac{M!}{a_1! a_2! \dots} v_1^{a_1} v_2^{a_2} \dots, \\ = (v_1 + v_2 + \dots)^M = V^M,$$

by the multinomial theorem. This could have been deduced at once by working direct with probabilities v_r/V instead of weights v_r , but the argument has been given in detail to illustrate the method for more complicated cases.

We next find the average value of a_r . This is given by

$$C\bar{a}_r = \sum_a a_r \frac{M!}{a_1! a_2! \dots} v_1^{a_1} v_2^{a_2} \dots.$$

To sum this expression we only have to cancel a_r with the first factor in $a_r!$ in the denominator, and then it is seen to be equal to

$$M v_r (v_1 + v_2 + \dots)^{M-1},$$

and so, as is implicit in our assumptions,

$$\bar{a}_r = M v_r / V. \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.3)$$

But we can now go further and find the range over which a_r will be likely to fluctuate. This is estimated by averaging the square of the difference of a_r from its mean value. We shall throughout this paper describe such a mean square departure as the *fluctuation* of the corresponding quantity. Thus the *fluctuation* of a_r is $(a_r - \bar{a}_r)^2$. Now

$$(a_r - \bar{a}_r)^2 = a_r(a_r - 1) + a_r - 2a_r \bar{a}_r + \bar{a}_r^2,$$

and averaging the separate terms by the multinomial theorem, we have

$$\overline{(a_r - \bar{a}_r)^2} = \frac{M(M-1)v_r^2}{V^2} + \frac{M v_r}{V} - \frac{2M v_r}{V} \frac{M v_r}{V} + \left(\frac{M v_r}{V}\right)^2 \\ = \frac{M v_r}{V} \left(1 - \frac{v_r}{V}\right) = \bar{a}_r \left(1 - \frac{\bar{a}_r}{M}\right). \quad . \quad . \quad . \quad (3.4)$$

This result represents the *fluctuation* however large or small

v_r may be. In all cases we have the result that $(\overline{a_r} - \overline{a_r})^2$ is less than $\overline{a_r}$, and therefore that the average departure of a_r from $\overline{a_r}$ is of order $(\overline{a_r})^{\frac{1}{2}}$. We can also interpret this fact by saying that departures of a_r from $\overline{a_r}$ which are much greater than $(\overline{a_r})^{\frac{1}{2}}$ will be relatively rare; as M is large and $(\overline{a_r})^{\frac{1}{2}}$ small compared with a_r , this is precisely equivalent to saying that the possession of the average value of a_r is a *normal* property of the assembly in the sense used by Jeans*. We have thus a simple and *complete* proof that uniform density is a normal property of this assembly.

§ 4. *The Distribution of Energy among a Set of Planck Vibrators.*

Another case where the treatment can be almost entirely algebraic is that of the partition of energy among a set of Planck vibrators which all have the same frequency. Let ϵ be the unit of energy so that every vibrator can have any multiple of ϵ . As we saw in § 2, the weight attached to every state is to be taken as unity.

Let there be M vibrators and let there be Pe of energy (P is an integer) to be partitioned among them. To specify a statistical state, let a_0 be the number of vibrators with no energy, a_1 with ϵ , a_2 with 2ϵ , etc. Then we have

$$a_0 + a_1 + a_2 + a_3 + \dots = M, \quad . \quad . \quad . \quad (4.1)$$

$$a_1 + 2a_2 + 3a_3 + \dots = P, \quad . \quad . \quad . \quad (4.2)$$

and any set of a 's which satisfy these equations corresponds to a possible state of the assembly. By the principles of § 2 each of the complexions will have unit weight. Now count up the number of complexions corresponding to the specification. By considering the various permutations of the vibrators, it is seen to be

$$\frac{M!}{a_0! a_1! a_2! \dots} \quad . \quad . \quad . \quad (4.3)$$

We must next find C the total number of all possible complexions. Let Σ_a denote summation over all possible values of the a 's which satisfy (4.1) and (4.2). Then

$$C = \Sigma_a \frac{M!}{a_0! a_1! a_2! \dots}$$

Consider the infinite series

$$(1 + z + z^2 + z^3 + \dots)^M$$

* Jeans, 'Dynamical Theory of Gases,' *passim*.

expanded by the multinominal theorem. The typical term is

$$\frac{M!}{a_0! a_1! a_2! \dots} z^{a_1+2a_2+3a_3 \dots},$$

where the a 's take any values consistent with (4.1). Then if we pick out the coefficient of z^P , we have the sum of all the expressions for which the a 's satisfy both (4.1) and (4.2). Observe that we may take the whole infinite series because the later terms are automatically excluded.

Now this will be the coefficient of z^P in $(1-z)^{-M}$, and so

$$C = \frac{(M+P-1)!}{(M-1)! P!}, \dots \dots \dots (4.4)$$

which is the ordinary expression for the number of homogeneous products as formerly used by Planck*.

We next evaluate the average of a_r ;

$$\begin{aligned} \bar{C}a_r &= \sum_a a_r \frac{M!}{a_0! a_1! a_2! \dots}, \\ &= M \sum_{a'} \frac{(M-1)!}{a_0'! a_1'! a_2'! \dots}, \end{aligned}$$

where $\sum_{a'}$ denotes summation over all values satisfying

$$\begin{aligned} a_0' + a_1' + a_2' + a_3' \dots &= M-1, \\ a_1' + 2a_2' + 3a_3' \dots &= P-r. \end{aligned}$$

The sum is thus

$$M \frac{(M+P-r-2)!}{(M-2)!(P-r)!},$$

and we have

$$\bar{a}_r = M(M-1) \frac{(M+P-r-2)!}{(M+P-1)!} \frac{P!}{(P-r)!}.$$

This is exact, and holds for all values of r ; now r can have any value up to P and the majority of the a_r 's will be zero. The ordinary method of proof applies Stirling's formula for $a_r!$ to these zero values. In the important case where both M and P are large, it will be only necessary to consider values of r which are small compared with P . Now, if r^2 is small compared with P , $P!/(P-r)!$ has the asymptotic value P^r . Using relations of this type and also disregarding the difference between M and $M-1$, we have

$$\bar{a}_r = \frac{M^2 P^r}{(M+P)^{r+1}} \dots \dots \dots (4.5)$$

* M. Planck in the earlier editions of his book on Radiation.

The same methods give the fluctuations of a_r . For

$$\overline{(a_r - \bar{a}_r)^2} = \overline{a_r a_r - 1} + \bar{a}_r - (\bar{a}_r)^2,$$

and a process similar to the above gives

$$\overline{C a_r (a_r - 1)} = M(M-1) \frac{(M+P-2r-3)!}{(M-3)!(P-2r)!}.$$

The exact expression for the fluctuation can be at once put down. When M and P are taken large the leading term cuts out, and so it is necessary to carry the approximation to the next order. If we substitute

$$P!/(P-r)! \sim P^{r-\frac{1}{2}} r(r-1) P^{r-1},$$

we find that

$$\overline{(a_r - \bar{a}_r)^2} = \bar{a}_r \left\{ 1 - \frac{M P^{r-1}}{(\bar{M} + \bar{P})^{r+2}} [2P^2 - (2r-1)MP + r^2 M^2] \right\}. \quad (4.6)$$

The formula for \bar{a}_r can be put into a more familiar form by the substitution $P = M/(e^\alpha - 1)$, which gives

$$\bar{a}_r = M e^{-r\alpha} (1 - e^{-\alpha}), \quad . \quad . \quad . \quad (4.7)$$

and leads to a corresponding but more complicated expression for the fluctuation. Here, as we shall see later, α can be identified with the familiar ϵ/kT . Equation (4.6) establishes at once that the statistical state specified by (4.5) or (4.7) is a normal property of the assembly.

§ 5. *The Partition of Energy among two Sets of Planck Vibrators.*

After these preliminary examples we now apply our method to a problem which will bring out its distinctive character, that of the partition of energy in an assembly composed of different types of system. We shall consider first the simplest of such cases—an assembly consisting of a large number of Planck vibrators of two types A and B. The number of A's is M , and the energy unit of an A is ϵ as before. There are now also N B's with energy unit η . To make exchanges of energy possible we have to suppose, say, that there are present a few gas molecules, but that the latter never possess any sensible amount of energy. (Later on in § 12 we develop a method by which we shall be able to include any number of such molecules in our assembly.) We also require for the purposes of the proof to assume that ϵ and η are commensurable, but it does not matter how large the numbers may be which are required to express the ratio ϵ/η in its

lowest terms. To avoid introducing new symbols, we may suppose that the unit of energy is so chosen that ϵ and η are themselves integers without a common factor.

We have already introduced the idea of weight, and seen that we must assign the weight unity to every permissible state of a linear vibrator. To calculate the number of complexions of the assembly of any given sort, we have merely to calculate the number of ways in which the energy may be distributed among the vibrators, subject to the given statistical specification. A simple example will make the process clear.

Let there be two A's and two B's; let $\eta = 2\epsilon$, $E = 4\epsilon$. Then the possible complexions are:—

$$\begin{array}{ccc} \left\{ \begin{array}{l} aaaa \\ a'a'a'a' \end{array} \right. & \left\{ \begin{array}{l} aab \\ aab' \\ a'a'b \\ a'a'b' \end{array} \right. & \left\{ \begin{array}{l} bb \\ b'b' \end{array} \right. \\ \left\{ \begin{array}{l} aaaa' \\ aa'a'a' \end{array} \right. & \left\{ \begin{array}{l} aa'b \\ aa'b' \end{array} \right. & \\ aaa'a' & & bb' \end{array}$$

Here, for example, $aaaa'$ means that there is 3ϵ of energy on the first of the A's, ϵ on the second, and none on the B's. Each of the fourteen complexions is, by definition, of equal weight, and is therefore to be reckoned as of equal probability in the calculation of averages. Observe how with the small amount of energy available a good deal more goes into the smaller than into the larger quanta; for the pair of A's have on the average $\frac{16}{7}\epsilon$, as against $\frac{12}{7}\epsilon$ for the pair of B's.

We pass to the general case. The statistical state of the assembly is specified by sets of numbers a_r, b_s , where a_r is the number of vibrators of type A which have energy $r\epsilon$, and b_s the number of B's with energy $s\eta$. All weights are unity and the number of complexions representing this statistical state is the number of indistinguishable ways (combinations) in which M vibrators can be divided into sets of $a_0, a_1 \dots$ and at the same time N into sets $b_0, b_1 \dots$. As illustrated by the example, it is therefore given by the formula

$$\frac{M!}{a_0! a_1! a_2! \dots} \frac{N!}{b_0! b_1! b_2! \dots} \dots \quad (5.1)$$

In (5.1) a_r and b_s may have any zero or positive values consistent with the conditions

$$\sum_r a_r = M, \quad \sum_s b_s = N, \quad \sum_r r\epsilon a_r + \sum_s s\eta b_s = E, \quad (5.2)$$

where E is the total energy of the system—necessarily an

integer in the units we employ. The total number C of all complexions is therefore

$$C = \sum_{a,b} \frac{M!}{a_0! a_1! a_2! \dots} \frac{N!}{b_0! b_1! b_2! \dots} \dots \quad (5.3)$$

where the summation $\sum_{a,b}$ is to be carried out over all positive or zero values of a_r and b_s which satisfy (5.2).

By using (5.1) and (5.3) we can at once obtain an expression for the average value, taken over all complexions, of any quantity in which we are interested. We have already studied \bar{a}_r in § 4. The main interest centres in \bar{E}_A , the average energy on the A 's. We have at once

$$C\bar{E}_A = \sum_{a,b} \frac{(\sum_r r \epsilon a_r) M! N!}{a_0! a_1! a_2! \dots b_0! b_1! b_2! \dots} \dots \quad (5.4)$$

The following process leads to simple integrals to express the quantities C , $C\bar{E}_A$, etc. Consider the infinite series

$$(1 + z^\epsilon + z^{2\epsilon} + \dots)^M$$

expanded in powers of z by the multinomial theorem. The general term is

$$\frac{M!}{a_0! a_1! \dots} z^{\sum_r r \epsilon a_r}.$$

It follows that if we select from the expansion of

$$(1 + z^\epsilon + z^{2\epsilon} + \dots)^M (1 + z^\eta + z^{2\eta} + \dots)^N \dots \quad (5.5)$$

the coefficient of z^E , we shall obtain the sum of all possible terms such as (5.1) subject to the conditions (5.2), that is to say C . Similarly, if we form the expression

$$\left\{ z \frac{d}{dz} (1 + z^\epsilon + z^{2\epsilon} \dots)^M \right\} (1 + z^\eta + z^{2\eta} \dots)^N, \dots \quad (5.6)$$

the general term in the first bracket must be

$$\frac{(\sum_r r \epsilon a_r) M!}{a_0! a_1! \dots} z^{\sum_r r \epsilon a_r},$$

and by the same reasoning the coefficient of z^E in (5.6) must be $C\bar{E}_A$.

Expressions (5.5) and (5.6) are easily simplified—they are respectively

$$(1 - z^\epsilon)^{-M} (1 - z^\eta)^{-N}, \quad \left\{ z \frac{d}{dz} (1 - z^\epsilon)^{-M} \right\} (1 - z^\eta)^{-N}.$$

The latter can also be written as

$$\left\{ -Mz \frac{d}{dz} \log(1-z^\epsilon) \right\} (1-z^\epsilon)^{-M} (1-z^\eta)^{-N}.$$

If these expressions are now expanded in powers of z by the binomial theorem, they give a sum of products of factorials which are, of course, the "homogeneous product" expressions used by Planck. It is possible to approximate to these by a legitimate use of Stirling's theorem and to replace the sums by integrals without much difficulty. It would, indeed, have been possible to start from these expressions, but we have not done so because in the general case to be discussed later that method would not be available. To make further progress by a method of general utility, we discard Stirling's theorem and express these coefficients of z^E by contour integrals taken round a circle γ with centre at the origin and radius less than unity. By well-known theorems on integration* we have at once

$$C = \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z^{E+1}} \frac{1}{(1-z^\epsilon)^M (1-z^\eta)^N}, \quad \dots \quad (5.71)$$

$$C \bar{E}_A = \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z^{E+1}} \frac{-Mz \frac{d}{dz} \log(1-z^\epsilon)}{(1-z^\epsilon)^M (1-z^\eta)^N} \dots \quad (5.72)$$

We can no longer hope for the single-term formulæ of §§ 3, 4. But (5.71) (5.72) are *exact*, and when M, N, E are all large in any definite fixed ratios, we can make use of the *method of steepest descents* to obtain simple adequate approximations. The method is very powerful and can be applied in a great number of cases without difficulty. Moreover, it is comparatively easy to use it with mathematical rigour if that is desired; and thus the somewhat clumsy calculations in the usual proofs of partition theorems are entirely avoided.

In general terms the process is this. Consider the integrand on the real positive axis. It becomes infinite at $z=0$ and again at $z=1$, and somewhere between at $z=\delta$ there is a minimum which is easily shown to be unique. Take as the contour the circle with centre at the origin and radius δ passing through this minimum. Then we find that, for

* For those not familiar with these theorems we may remark that $\frac{1}{2\pi i} \int_{\gamma} z^r dz = 0$ when r is any integer other than -1 , while $\frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z} = 1$; these equations at once give (5.71) and (5.72).

values of z on the contour, $z = \mathfrak{S}$ corresponds to a strong maximum, and when M, N, E are large, such a strong maximum that practically the whole value of the integral is contributed by the contour in the neighbourhood of this point. Hence in the integrals it is legitimate to substitute the value at this point for any factors which do not themselves show strong maxima here or elsewhere. On this general principle we can remove the term $-Mz \frac{d}{dz} \log(1-z^\epsilon)$ from under the integral sign, provided that z is given the value \mathfrak{S} determined by the maximum condition. The part of the integrand involving the large numbers M, N, E , determines the value of \mathfrak{S} as being the unique real positive fractional root of the equation

$$\frac{d}{dz} \left\{ z^{-E}(1-z^\epsilon)^{-M}(1-z^\eta)^{-N} \right\} = 0.$$

That is, \mathfrak{S} satisfies the equation :

$$E = \frac{M\epsilon}{\mathfrak{S}^{-\epsilon}-1} + \frac{N\eta}{\mathfrak{S}^{-\eta}-1} \cdot \cdot \cdot \cdot \cdot \quad (5.8)$$

The remaining integrands in C and $C\overline{E_A}$ are identical, and we therefore have

$$\begin{aligned} \overline{E_A} &= -M\mathfrak{S} \frac{d}{d\mathfrak{S}} \log(1-\mathfrak{S}^\epsilon), \\ &= \frac{M\epsilon}{\mathfrak{S}^{-\epsilon}-1} \cdot \cdot \cdot \cdot \cdot \quad (5.9) \end{aligned}$$

If a similar process is carried out for the B 's, we have

$$\overline{E_B} = \frac{N\eta}{\mathfrak{S}^{-\eta}-1}, \cdot \cdot \cdot \cdot \cdot \quad (5.91)$$

in agreement with the necessary relation

$$\overline{E_A} + \overline{E_B} = E.$$

Equations (5.9) (5.91) determine the partition of energy and take their familiar form if we replace \mathfrak{S} by $e^{-1/KT}$. We shall return to this point later.

§ 6. Application of the Method of Steepest Descents.

After this sketch it will be well to establish the validity of the arguments used. This section is put in for mathematical completeness, and is not concerned at all with physical questions. We treat of a more general case than that of § 5.

Arguments of this type—asymptotic expansions by steepest descents—are, of course, well known in pure mathematics.

Consider a contour integral of the form

$$\frac{1}{2\pi i} \int_{\gamma} F(z) [\phi(z)]^E \frac{dz}{z}, \quad \dots \quad (6.1)$$

subject to the following conditions:—

- (i.) $\phi(z)$ is an analytic function of z , which can be expanded in a series of ascending powers of z .
- (ii.) This series starts with some negative power.
- (iii.) Every coefficient is real and positive.
- (iv.) Its circle of convergence is of radius unity. (This condition is quite unessential to the mathematics, but makes the statement simpler, and is physically true.)
- (v.) The powers that occur in the series cannot all be put in the form $\alpha + \beta r$ where α and β are any given integers and r takes all integer values.
- (vi.) $F(z)$ is an analytic function with no poles in the unit circle, except perhaps at the origin.
- (vii.) γ is a closed contour going once counterclockwise round the origin.

The problem is to obtain the asymptotic value as E tends through integer values to infinity.

We shall first study the properties of $\phi(z)$. Considering real values, it must have one and only one minimum between 0 and 1. For it is continuous and tends to $+\infty$ at both 0 and 1, and so must have at least one minimum between. Further, to find minima we differentiate, and then all the negative powers will have negative coefficients and all the positive positive. It follows that minima are given where two curves cut, one of which decreases steadily between 0 and 1, while the other increases steadily. These curves can only cut in one point, and so there is only one minimum. Next, for the complex values, consider a circle of any radius r less than 1. As the modulus of a sum is never greater than the sum of the moduli, it follows that at no point on this circle can $|\phi(z)|$ be greater than $\phi(r)$. Moreover, it can only equal $\phi(r)$ provided that condition (v.) is broken, and in that case there will be β points at each of which $|\phi(z)| = \phi(r)$. We can thus see that on account of the large exponent it is only the part of γ near the real axis that contributes effectively to the asymptotic value of the integral. This suggests

the substitution $z = re^{i\alpha}$, with α as the new variable of integration. Expanding near the real axis we have

$$[\phi(z)]^E = [\phi(r)]^E \exp. \{i\alpha E r \phi' / \phi + O(E\alpha^2)\}. \quad (6.11)$$

This function contains a periodic term of high frequency which cuts down the contributions for small values of α , so that the value of the integral will in general depend on more distant parts of the contour than those to which this approximation will apply. But if we choose for r the special value \mathfrak{S} corresponding to the unique minimum $\phi' = 0$, then the oscillating term in the exponential vanishes and the contributions for small values of α dominate the whole integral. For this special value of r the exponential becomes

$$\exp. \left\{ -\frac{1}{2} E \alpha^2 \mathfrak{S}^2 \phi'' / \phi + O(E\alpha^3) \right\}, \quad . \quad . \quad (6.12)$$

and by (iii.) $\phi'' > 0$. We see at once that we can suppose that $E^{\frac{1}{2}}\alpha$ ranges effectively over all values from $-\infty$ to $+\infty$ while all other terms, such as $\alpha, \dots, E\alpha^3, \dots$ remain small. We then obtain for (6.1) on putting $z = \mathfrak{S}e^{i\alpha}$ the asymptotic expression

$$\frac{1}{2\pi} [\phi(\mathfrak{S})]^E \int_{-\infty}^{+\infty} \{F(\mathfrak{S}) + i\alpha \mathfrak{S} F'(\mathfrak{S}) + O(\alpha^2) + O(E\alpha^3)\} e^{-\frac{1}{2} E \alpha^2 \mathfrak{S}^2 \phi'' / \phi} d\alpha.$$

For most purposes the first term in the expansion will suffice, but if the precise values of the fluctuations are required, the second also is necessary. As it is in general rather complicated, we shall content ourselves here with pointing out its order of magnitude.

On carrying through the necessary calculations we find

$$\frac{\phi^E}{(2\pi E \mathfrak{S}^2 \phi'' / \phi)^{\frac{1}{2}}} \left[F - \frac{\{\mathfrak{S}, \phi, F\}}{2E \mathfrak{S}^2 \phi'' / \phi} \right]. \quad . \quad . \quad (6.2)$$

The argument of F and ϕ is everywhere \mathfrak{S} ; the term $\{\mathfrak{S}, \phi, F\}$ denotes a complicated expression of \mathfrak{S}, ϕ and its first four derivatives, and F and its first two derivatives, but is independent of E . If condition (v.) is dropped, we shall have β equal maxima arranged round the circle γ , and, provided F has the same value at each of them, the integral will have a value equal to (6.2) multiplied by β .

Now consider the problem of § 5, to which our work applies immediately with

$$\phi = z^{-1}(1 - z^e)^{-M/E}(1 - z^{\eta})^{-N/E}. \quad . \quad . \quad (6.3)$$

We may suppose that E tends to infinity and also M and N

in such a way that M/E and N/E are constant. This function satisfies all the conditions of this section—the fact that it is in general many-valued is irrelevant, for we are only concerned with that particular branch which is real when z is real and $0 < z < 1$ and this branch is one-valued in the unit circle. The unique minimum \mathfrak{S} is determined by the equation $\phi' = 0$ or

$$E = \frac{M\epsilon}{\mathfrak{S}^{-\epsilon}-1} + \frac{N\eta}{\mathfrak{S}^{-\eta}-1} \quad \dots \quad (6.4)$$

The value of the integral (5.71) is then by (6.2) (omitting the second term the form of which is only required in calculating fluctuations),

$$C = \frac{\mathfrak{S}^{-E}(1-\mathfrak{S}^\epsilon)^{-M}(1-\mathfrak{S}^\eta)^{-N}}{\{2\pi E \mathfrak{S}^2 \phi''/\phi\}^{\frac{1}{2}}} \quad \dots \quad (6.5)$$

If, contrary to hypothesis, we had taken ϵ and η as having a common factor β , condition (v.) would have been violated. In this case C would be β times as great as before, but so would $C\bar{E}_A$, so that \bar{E}_A and all other averages would be unaffected. The use of (6.2) to evaluate $C\bar{E}_A$ (5.72) etc. leads at once to expressions similar to (6.5), and so to the results given formally in the last section and to others to be given later.

As we shall see, \mathfrak{S} is the temperature measured on a special scale, and there is great advantage in regarding \mathfrak{S} , rather than E , as the independent variable which determines the state of the assembly. If this is done the expression $E\mathfrak{S}^2\phi''/\phi$ can be put into simpler form. For it is easily verified that

$$\begin{aligned} E\mathfrak{S}^2\phi''/\phi &= \frac{M\epsilon\mathfrak{S}^{-\epsilon}}{(\mathfrak{S}^{-\epsilon}-1)^2} + \frac{N\eta^2\mathfrak{S}^{-\eta}}{(\mathfrak{S}^{-\eta}-1)^2}, \\ &= \mathfrak{S} \frac{d}{d\mathfrak{S}} (\bar{E}_A + \bar{E}_B) = \mathfrak{S} \frac{dE}{d\mathfrak{S}}, \quad \dots \quad (6.6) \end{aligned}$$

if E is regarded as a function of \mathfrak{S} , given by (5.8).

It may be remarked that the constant occurrence of the operator $\mathfrak{S}d/d\mathfrak{S}$ suggests a change of variable to $\log \mathfrak{S}$. Though this has some advantages we have not adopted it, partly because it makes the initial argument about the multinomial theorem a little harder to follow and complicates the contour of the integration, and also because $\log \mathfrak{S}$ is not itself the absolute temperature—if it had been, the physical simplicity might have outweighed the other objections—but only a quantity proportional to its reciprocal.

§ 7. *The Meaning of \mathfrak{S} .*

Returning to the subject of § 5, we see that the partition of energy between the A's and the B's depends on a parameter which can be determined in terms of the energy by means of the equation (5.8). It is fairly evident that \mathfrak{S} is connected with the temperature, though it requires more general considerations to prove this properly. But if we assume this connexion and are content to replace the thermodynamic definition of absolute temperature by one based on the law of equipartition of energy for systems obeying the laws of classical mechanics, then we can at once identify the meaning of \mathfrak{S} .

For let us suppose that the B's are vibrators of very low frequency. They will then obey the classical laws, and the average energy of each will be kT . But

$$\lim_{\eta \rightarrow 0} \frac{\eta}{\mathfrak{S}^{-\eta} - 1} = \frac{1}{\log(1/\mathfrak{S})} = kT,$$

which shows that

$$\mathfrak{S} = e^{-1/kT}. \quad \dots \dots (7.1)$$

Substituting this in (5.9), we obtain the well-known form

$$\overline{E}_A = \frac{M\epsilon}{e^{\epsilon/kT} - 1}.$$

Observe that while η is tending to zero there need be no difficulty about the condition that η and ϵ are to be commensurable. We shall later return to this question of temperature and establish it for much more general types of system.

§ 8. *Other mean values.*

Exactly the same methods can be applied to evaluate any other mean value besides \overline{E}_A . For example :—

$$\begin{aligned} C\bar{a}_r &= \sum_{a,b} a_r \frac{M!}{a_0! a_1! \dots} \frac{N!}{b_0! b_1! \dots}, \\ &= M \sum_{a',b} \frac{(M-1)!}{a_0'! a_1'! \dots} \frac{N!}{b_0! b_1! \dots}, \end{aligned} \quad (8.1)$$

summed over all zero and positive values such that

$$\sum_r a_r' = M-1, \quad \sum_s b_s = N, \quad \sum_r r \epsilon a_r' + \sum_s s \eta b_s = E - r\epsilon.$$

Applying the multinomial theorem and reducing the

expression to a complex integral, we have

$$C\bar{a}_r = M \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z^{\frac{M}{E}+1}} \frac{z^{r\epsilon}}{(1-z^{\epsilon})^{M-1} (1-z^{\eta})^N}, \quad (8.2)$$

which, by virtue of the value of C from (5.71) and the argument of § 6, at once yields

$$\begin{aligned} \bar{a}_r &= M \mathfrak{S}^{r\epsilon} (1 - \mathfrak{S}^{\epsilon}), \\ &= M e^{-r\epsilon/kT} (1 - e^{-\epsilon/kT}), \quad \dots \quad (8.3) \end{aligned}$$

which is the formula of § 4 over again, the presence of the B 's being immaterial.

When we come to evaluate fluctuations the matter is a little more complicated, because the leading terms cut out, and so the second term of the asymptotic expansion will in general play a part. For example, consider the fluctuations of a_r :

$$(\overline{a_r - \bar{a}_r})^2 = \overline{a_r(a_r - 1)} + \bar{a}_r - (\bar{a}_r)^2.$$

By arguments exactly similar to those above, we have

$$C\overline{a_r(a_r - 1)} = M(M-1) \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z^{\frac{M}{E}+1}} \frac{z^{2r\epsilon}}{(1-z^{\epsilon})^{M-2} (1-z^{\eta})^N},$$

and so by (6.2)

$$\overline{a_r(a_r - 1)} = M(M-1) \mathfrak{S}^{2r\epsilon} (1 - \mathfrak{S}^{\epsilon})^2 \{1 + O(1/E)\}.$$

Thus the fluctuation is

$$\bar{a}_r - M \mathfrak{S}^{2r\epsilon} (1 - \mathfrak{S}^{\epsilon})^2 + O(M^2/E).$$

This is sufficient to show that the possession of \bar{a}_r (8.3) is a normal property of the assembly.

The complete calculation of the O -term is rather complicated; the result is given at the end of this section. But a great simplification arises if we suppose that there are many more B 's than A 's, while E is so adjusted that \mathfrak{S} the temperature is unchanged. In this case the term $O(M^2/E)$ becomes small and may be neglected. We shall describe this case by saying that A is in a bath of temperature \mathfrak{S} . Then, provided this is so, we have

$$(\overline{a_r - \bar{a}_r})^2 = \bar{a}_r (1 - \bar{a}_r/M). \quad \dots \quad (8.4)$$

A much more important quantity is the fluctuation of E_A . This is found by evaluating $\overline{E_A^2}$. Now, just as $C\bar{E}_A$ was given by operating with zd/dz on the first factor in

$(1-z^e)^{-M}(1-z^n)^{-N}$, so $C\overline{E_A^2}$ is easily seen to be given by operating with $\left(z\frac{d}{dz}\right)^2$ in the same way. Thus

$$C\overline{E_A^2} = \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z^{\overline{E}+1}} \left\{ \left(z\frac{d}{dz}\right)^2 (1-z^e)^{-M} \right\} (1-z^n)^{-N}.$$

If we again suppose an infinite bath of temperature \mathfrak{S} , we can omit the second term of the asymptotic expansion (6.2) and obtain

$$\begin{aligned} \overline{E_A^2} &= (1-\mathfrak{S}^e)^M \left\{ \left(\mathfrak{S}\frac{d}{d\mathfrak{S}}\right)^2 (1-\mathfrak{S}^e)^{-M} \right\}, \\ &= (1-\mathfrak{S}^e)^M \mathfrak{S} \frac{d}{d\mathfrak{S}} \left\{ \overline{E_A} (1-\mathfrak{S}^e)^{-M} \right\}, \\ &= (\overline{E_A})^2 + \mathfrak{S} \frac{d\overline{E_A}}{d\mathfrak{S}}, \end{aligned}$$

and so the fluctuation is

$$\left. \begin{aligned} (\overline{E_A - E_A})^2 &= \overline{E_A^2} - (\overline{E_A})^2, \\ &= \mathfrak{S} \frac{d\overline{E_A}}{d\mathfrak{S}}, \\ &= kT^2 \frac{d\overline{E_A}}{dT}. \end{aligned} \right\} \dots \dots (8.5)$$

This is a result of which Einstein * made use in his work on fluctuations of radiation. It should be emphasized that these results are only accurate in a temperature bath, and not when the number of systems A is a finite fraction of the assembly.

In all cases (6.2) shows that the possession of $\overline{E_A}$ is a normal property of the assembly.

If we work out exactly the second terms in the asymptotic formulæ of § 6 and apply them to the fluctuations of a_r and E_A we find

$$(\overline{a_r - a_r})^2 = \overline{a_r} \left[1 - \frac{\overline{a_r}}{M} \left\{ 1 + \frac{M(re - \overline{E_A}/M)^2}{\mathfrak{S} d\overline{E_A}/d\mathfrak{S}} \right\} \right], \quad (8.6)$$

$$(\overline{E_A - E_A})^2 = \mathfrak{S} \frac{d\overline{E_A}}{d\mathfrak{S}} \left(1 - \mathfrak{S} \frac{d\overline{E_A}}{d\mathfrak{S}} / \mathfrak{S} \frac{d\overline{E}}{d\mathfrak{S}} \right). \quad \dots \dots (8.7)$$

Formula (4.6) above is a special case of (8.6).

* A. Einstein, *Phys. Zeitschr.* vol. x. p. 185 (1909).

Finally it is of some interest to point out that we can obtain a formula for $\overline{(E_A - \bar{E}_A)^{2s}}$ of general validity. We have in fact

$$\overline{(E_A - \bar{E}_A)^{2s}} = 1.3 \dots (2s-1) \{ \overline{(E_A - \bar{E}_A)^2} \}^s, \quad (8.8)$$

where $\overline{(E_A - \bar{E}_A)^2}$ is given by (8.7). We retain of course only the highest order term *, which is thus $O(\bar{E}_A)^s$.

§ 9. *Generalization to any number of types of system, and to systems of any quantized character.*

It is clear that the present method of treating partitions is of a much more general character than has so far been exhibited. Consider an assembly composed of two types, A and B, of quantized systems more complicated than Planck vibrators. We suppose generally that the systems of type A, M in number, can take energies to the extents $\epsilon_0, \epsilon_1, \epsilon_2, \dots$, and these states have *weight* factors p_0, p_1, p_2, \dots in conformity with the discussion in § 2. Similarly, the B's, N in number, can take energies $\eta_0, \eta_1, \eta_2, \dots$ with weights q_0, q_1, q_2, \dots . We have to suppose that it is possible to determine a basal unit of energy such that all the ϵ 's and η 's can be expressed as integers. Further, it simplifies the work if we suppose that there is no factor common to all of them. Proceeding exactly as before, we set down the weighted number of complexions which correspond to the specification that, of the A's, a_r have energy ϵ_r ; of the B's, b_s have energy η_s . This number is

$$\frac{M!}{a_0! a_1! \dots} p_0^{a_0} p_1^{a_1} \dots \frac{N!}{b_0! b_1! \dots} q_0^{b_0} q_1^{b_1} \dots, \quad (9.1)$$

and the a 's and b 's are able to take all values consistent with

$$\sum_r a_r = M, \quad \sum_s b_s = N, \quad \sum_r \epsilon_r a_r + \sum_s \eta_s b_s = E. \quad (9.2)$$

Now form the functions

$$f(z) = p_0 z^{\epsilon_0} + p_1 z^{\epsilon_1} + p_2 z^{\epsilon_2} + \dots, \quad (9.3)$$

$$g(z) = q_0 z^{\eta_0} + q_1 z^{\eta_1} + q_2 z^{\eta_2} + \dots \quad (9.31)$$

These will be called the *partition functions* † of the types of

* Cf. Gibbs' 'Statistical Mechanics,' p. 78. But (8.8) is generally valid, while Gibbs' formulæ really refer only to a group of systems in a temperature bath.

† They are practically the "Zustandsumme" of Planck, 'Radiation Theory,' p. 127.

system A and B. The application of the multinomial theorem then leads to the consideration of the expression

$$[f(z)]^M [g(z)]^N,$$

and pursuing exactly the same course as in § 5, we find

$$C = \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z^{E+1}} [f(z)]^M [g(z)]^N, \quad \dots \quad (9.4)$$

$$C\bar{E}_A = \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z^{E+1}} \left\{ z \frac{d}{dz} [f(z)]^M \right\} [g(z)]^N. \quad (9.5)$$

Assume for the moment that we can choose a $\phi(z)$ conforming to the requirements of § 6. The whole calculation then goes on as before. The radius of the circle to be taken as contour is given by the equation

$$E = M\vartheta \frac{d}{d\vartheta} \log f(\vartheta) + N\vartheta \frac{d}{d\vartheta} \log g(\vartheta). \quad \dots \quad (9.6)$$

This equation has one and only one root. We thus can at once put down

$$\begin{aligned} \bar{E}_A &= [f(\vartheta)]^{-M} \vartheta \frac{d}{d\vartheta} [f(\vartheta)]^M \\ &= M\vartheta \frac{d}{d\vartheta} \log f(\vartheta). \end{aligned} \quad \dots \quad (9.7)$$

In exactly the same way we have

$$\bar{a}_r = M p_r \vartheta^{e_r} / f(\vartheta), \quad \dots \quad (9.8)$$

and we can also verify that in the case of an infinite bath the fluctuations are again given by (8.4), (8.5), and that equation (6.6) is still true. The exact forms of the fluctuations (8.6), (8.7) are also valid if we replace ϵ_r by ϵ_r .

We have now to examine whether $\phi(z)$ can be properly chosen. It is natural to take

$$\phi(z) = z^{-1} [f(z)]^{M/E} [g(z)]^{N/E}. \quad \dots \quad (9.81)$$

By its definition it must satisfy (i.). For (ii.) to be true, we must have

$$E > M\epsilon_0 + N\eta_0,$$

which is the trivial condition that there must be enough energy to provide each system with the least amount of energy it is permitted to have. Condition (iv.) does not appear at first sight inevitable, but must follow from Bohr's Correspondence Principle*, for the convergence of the series $f(z)$ and $g(z)$ depends on their later terms—that is, those of

* Bohr, *loc. cit.*

large quantum numbers. Condition (v.) is satisfied if not all the ϵ 's and η 's have a common factor. There remains (iii.), and here there are trivial analytical difficulties when, as in general, M/E and N/E are fractional.

It is, however, easy to generalize § 6 by replacing $[\phi(z)]^E$ by

$$z^{-E}[f(z)]^M[g(z)]^N,$$

and letting E , M , N all tend to infinity independently. Condition (iii.) is then satisfied, as can be seen by multiplying out, and so all the conditions are satisfied, and the final results stated above are unaffected.

Finally, we may observe that all our results can be extended at once to an assembly containing any number of types of system. If there are M_c systems of type C , for which the partition function is $f_c(\mathfrak{S})$, then

$$\overline{E}_c = M_c \mathfrak{S} \frac{d}{d\mathfrak{S}} \log f_c(\mathfrak{S}),$$

where \mathfrak{S} is determined by

$$E = \sum_c M_c \mathfrak{S} \frac{d}{d\mathfrak{S}} \log f_c(\mathfrak{S}).$$

The formal validity of the proof will require all the quantities ϵ_c to be commensurable. It will be shown in § 12 how this restriction may be removed.

§ 10. Vibrators of two and three degrees of freedom.

As a first example we take a set of vibrators each of which is free to vibrate in a plane under a central force proportional to the distance. The sequence of energies is again $0, \epsilon, 2\epsilon, \dots$, but the weights are no longer unity, as the system is degenerate. Following the principle laid down in § 2, we may evaluate the weights by treating the system as non-degenerate and counting the number of different motions which have the same total quantum number. Now we can quantize the plane vibrators in directions x and y , and as an example for the case 2ϵ , we have three alternatives $(2\epsilon, 0)$, (ϵ, ϵ) , $(0, 2\epsilon)$. This is easily generalized, and gives to $r\epsilon$ the weight $r+1$. The partition function for these vibrators is thus

$$\begin{aligned} f(z) &= 1 + 2z + 3z^2 + 4z^3 + \dots, \\ &= (1-z)^{-2}. \end{aligned}$$

From the general theorem (9.7) we at once have

$$\overline{E}_A = \frac{2M\epsilon}{9^{-\epsilon} - 1},$$

so that such vibrators have just twice as much energy as the line vibrators.

In exactly the same way we can treat the case of three dimensions. To illustrate the weights we again take the case of 2ϵ and quantize the system in x, y, z . There are six alternatives $(2\epsilon, 0, 0)$, $(0, 2\epsilon, 0)$, $(0, 0, 2\epsilon)$, $(\epsilon, \epsilon, 0)$, $(\epsilon, 0, \epsilon)$, $(0, \epsilon, \epsilon)$. The general form for $r\epsilon$ is $\frac{1}{2}(r+1)(r+2)$. The partition function is now

$$\begin{aligned} f(z) &= 1 + 3z + 6z^2 + 10z^3 + \dots, \\ &= (1-z)^{-3}, \end{aligned}$$

which leads at once to the expected result

$$\overline{E}_A = \frac{3M\epsilon}{9^{-\epsilon} - 1}$$

§ 11. *Rotating Molecules.*

Another interesting example to which the calculations at once apply is that part of the specific heat of a gas due to the rotations of the molecules. Various writers* have quantized the motions of a rigid body, and it is found that the system has at most two instead of three periods, so that it is partly degenerate. We may consider for simplicity a diatomic molecule. Then, on account of the small moment of inertia about the line of centres, the third degree of freedom may be omitted altogether—its quantum of energy is too large. A simple calculation then leads to energies of rotation ϵ_r given by

$$\epsilon_r = \frac{h^2}{8\pi^2 I} r^2, \quad . \quad . \quad . \quad . \quad (11.1)$$

where I is the moment of inertia about a transverse axis, which we shall assume to be independent of r . This is a degenerate system, and considerations of the number of cases which occur if it is quantized for the two degrees shows that the weight to be attached is $2r+1$. This is on the principles suggested by Bohr † with a simplifying modification; for Bohr had to suppose that certain quantized motions were excluded for other reasons which are not operative

* Among others, Ehrenfest, *Verh. Deutsch. Phys. Ges.* xv. p. 451 (1913). Epstein, *Phys. Zeitschr.* xx. p. 289 (1919). F. Reiche, *Ann. der Physik*, liv. p. 421 (1917).

† *Loc. cit.* p. 26.

here *. There can, we think, be no question as to the correctness of the weight $2r+1$, but most recent writers have used a factor r ; our formula for the specific heat has therefore a rather different value.

We may now apply our general formulæ to this case with

$$p_r = 2r + 1, \quad \epsilon_r = r^2 \epsilon, \quad \epsilon = \frac{h^2}{8\pi^2 I}. \quad (11.2)$$

Then
$$f(\mathfrak{S}) = 1 + 3\mathfrak{S}^\epsilon + 5\mathfrak{S}^{4\epsilon} + 7\mathfrak{S}^{9\epsilon} + \dots, \quad (11.3)$$

$$E_A = M\mathfrak{S} \frac{d}{d\mathfrak{S}} \log f(\mathfrak{S}). \quad (11.31)$$

The contribution of the rotations to the molecular specific heat, C_{rot} , is dE_A/dT , where M must be taken as the number of molecules in one gramme-molecule of gas. Thus, using (7.1), we have

$$C_{rot} = \frac{Mk}{(\log \mathfrak{S})^2} \left(\mathfrak{S} \frac{d}{d\mathfrak{S}} \right)^2 \log f(\mathfrak{S}), \quad (11.4)$$

and $Mk = R$, the usual gas constant. If we write

$$\sigma = \frac{h^2}{8\pi^2 I k T}, \quad (11.5)$$

then

$$C_{rot} = R\sigma^2 \frac{d^2}{d\sigma^2} \log (1 + 3e^{-\sigma} + 5e^{-4\sigma} + 7e^{-9\sigma} + \dots). \quad (11.51)$$

Equation (11.5) shows that, when $T \rightarrow \infty$, $\sigma \rightarrow 0$. It can be shown by the application of standard theorems on series † that when $\sigma \rightarrow 0$,

$$C_{rot} \rightarrow R, \quad (11.52)$$

which is the correct limiting value as required by classical dynamics.

In the general case of any body we have three degrees of rotational freedom, the motion is simply degenerate ‡, and the energy enters as a sum of square numbers multiplying two units of energy. The motion of the axis of symmetry and the motion about the axis of symmetry are not independent, and it is impossible therefore for the partition function to split up into the product of two partition functions which represent the separate contributions of the two motions. The result is a *double* series of the same general type as (11.3).

* Assuming that no extraneous considerations rule out any of these states.

† Bromwich, *Infinite Series*, p. 132. The theorem is due to Cesàro.

‡ Epstein, *Phys. Zeit.* xx. p. 289 (1919).

It does not appear profitable to examine these expressions further here, since the agreement with experiment is not very good at all temperatures. It is to be presumed that the assumption of constant moments of inertia is at fault, and this is supported by some of the evidence from band spectra; further, it is probable that the case of no rotation must be excluded, involving the omission of the first term in the partition functions. The discussion of the practical applications of these formulæ cannot be entered into here.

§ 12. *Assemblies containing free molecules.*

The problems we have so far discussed have all possessed the distinguishing characteristic that the temperature is the only independent variable. As soon as we treat of free molecules this is no longer the case, for now the volume must be another independent variable. Nevertheless, as we shall see, the same methods of calculation are available. The partition is no longer represented exactly by integrals, as it was for the quantized motions, but from the nature of the case some form of limiting process is required. The free molecules cannot of course be regarded as the limit of three-dimensional vibrators of low frequency, for they have no potential energy to share in the partition. We must proceed by the method common to most discussions of the distribution laws of classical assemblies—divide up into cells the six-dimensional space in which the state of any molecule is represented, associate with each cell a certain constant value of the energy, and in the limit make all the dimensions of all the cells tend to zero*.

We take an assembly composed of M systems of the type A of § 9 and P free-moving monatomic molecules of mass m and of small size, the whole enclosed in a vessel of volume V . The energy of the molecules is solely their energy of translation; they are supposed to obey the laws of classical mechanics (except during their collisions with the A's). In order to specify the state of the assembly, we take a six-dimensional space of co-ordinates q_1, q_2, \dots, p_3 , the three rectangular co-ordinates and momenta of a molecule in the vessel. We divide up this space into small cells, 1, 2, 3, ..., t ..., of extensions $(dq_1 \dots dp_3)_t$, which may or may not be

* That the limit of the distribution laws worked out for the cells is the true distribution law for the actual assembly is an *assumption* implicit in all such discussions.

equal. Then by the principles of § 2 the weight factor for the t th cell is

$$\delta_t = \frac{(dq_1 \dots dp_s)_t}{h^3}, \quad . \quad . \quad . \quad (12.1)$$

provided of course that the cell is relevant to our assembly. Only those cells have a weight for which the q 's lie in the vessel; but the p 's may range over all values from $-\infty$ to $+\infty$, for the method of summation will automatically exclude values which could not be allowed. Associated with the t th cell there is energy given by

$$\zeta_t = \frac{1}{2m} (p_1^2 + p_2^2 + p_3^2). \quad . \quad . \quad . \quad (12.11)$$

The state of the molecules in the assembly is specified by the numbers c_1, c_2, \dots of molecules in cells 1, 2, The specification of the A 's is as before. The number of weighted complexions corresponding to the specification is then

$$\frac{M!}{a_0! a_1! \dots} p_0^{a_0} p_1^{a_1} \dots \frac{P!}{c_1! c_2! \dots} \delta_1^{c_1} \delta_2^{c_2} \dots, \quad (12.12)$$

$$\text{where} \quad \sum_r a_r = M, \quad \sum_t c_t = P, \quad \sum_r a_r \epsilon_r + \sum_t c_t \zeta_t = E. \quad (12.13)$$

In proceeding thus we are constructing an artificial assembly in which the energy is taken to have the same value ζ_t in all parts of the t th cell, and in which all the ζ 's and all the ϵ 's can be expressed as multiples of some basal unit, without a factor common to them all.

This assembly can be made to resemble the real one to any standard of approximation required. For such an artificial assembly we can make use of the whole of our machinery. The results all depend on integrals such as

$$C = \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z^{E+1}} [f(z)]^M [h(z)]^P, \quad . \quad . \quad (12.2)$$

where the partition function of the artificial molecules is

$$h(z) = \sum_t \delta_t z^{\zeta_t}, \quad . \quad . \quad . \quad (12.21)$$

and the formulæ of § 8 follow at once for $\overline{E_C}$, $(\overline{E_C} - \overline{E_C})^2$, $\overline{c_t}$ and $(\overline{c_t} - \overline{c_t})^2$. These results give completely the exact partition laws for any artificial assembly of the type considered. To obtain the actual distribution law for the real assembly, we must make all the dimensions of all the cells tend to zero, and obtain the limit of the partition function. Now, by the

476 Messrs. C. G. Darwin and R. H. Fowler on
definition of an integral, in the limit $\hbar(\mathfrak{S}) \rightarrow H(\mathfrak{S})$, where

$$H(\mathfrak{S}) = \frac{1}{h^3} \int^{(6)} e^{-\frac{\log 1/\theta}{2m}(p_1^2 + p_2^2 + p_3^2)} dq_1 \dots dp_3. \quad (12.3)$$

The integration is over the volume V and over all values of the p 's from $-\infty$ to $+\infty$. This gives at once

$$H(\mathfrak{S}) = \frac{(2\pi m)^{3/2} V}{h^3 (\log 1/\mathfrak{S})^{3/2}} \dots \quad (12.4)$$

In the formulæ the functions $dh/d\mathfrak{S}$ and $d^2h/d\mathfrak{S}^2$ also occur, and it is easily shown directly that their limits are $dH/d\mathfrak{S}$ and $d^2H/d\mathfrak{S}^2$. We may therefore use (12.4) throughout for the real assembly, and at once obtain the following expressions:

$$\overline{E}_C = \frac{\frac{3}{2}P}{\log 1/\mathfrak{S}} = \frac{3}{2}PkT, \quad \dots \quad (12.5)$$

$$c_i = \frac{P}{V} \left(\frac{m \log 1/\mathfrak{S}}{2\pi} \right)^{3/2} e^{-\frac{1}{2}(\log 1/\theta)m(u^2 + v^2 + w^2)} dx dy dz du dv dw, \quad \dots \quad (12.51)$$

$$= \frac{P}{V} \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-\frac{m}{2kT}(u^2 + v^2 + w^2)} dx \dots dw. \quad \dots \quad (12.52)$$

The temperature \mathfrak{S} is determined by

$$E = M\mathfrak{S} \frac{d}{d\mathfrak{S}} \log f(\mathfrak{S}) + \frac{3}{2}P \frac{1}{\log (1/\mathfrak{S})}, \quad \dots \quad (12.53)$$

and the fluctuation of energy of the molecules in a bath of temperature \mathfrak{S} is given by

$$(\overline{E}_C - \overline{E}_C)^2 = \mathfrak{S} \frac{d}{d\mathfrak{S}} \overline{E}_C = \frac{3}{2}Pk^2T^2 \quad \dots \quad (12.54)$$

If the fluctuation of c_i is evaluated, it takes the simple value

$$(\overline{c_i} - \overline{c_i})^2 = \overline{c_i}, \quad \dots \quad (12.55)$$

whether it is in a bath or not; for the second factor, analogous to that in (8.4), can be omitted when the cell is taken to be of small size. Thus in all cases the possession of $\overline{c_i}$ is a normal property of the assembly. These results can be readily extended to cases where there is an external field of force acting on the molecules.

By means of this assembly we can establish the meaning of \mathfrak{S} in terms of T , by observing that the gas itself constitutes a constant volume gas thermometer. It is easy to show that the pressure of a gas must be $\frac{2}{3}$ of the mean kinetic energy

in unit volume, that is to say, $p = P/V \log(1/\mathfrak{D})$. Since the gas temperature is measured by the relation $pV = PkT$, we are again led to the relation $\mathfrak{D} = e^{-1/kT}$.

We may observe that it is now possible to drop the assumption of commensurability, which was necessary in the sections which dealt with quantized systems. It was there essential, physically speaking, in order that it should be possible that the whole of the energy should be held somewhere; but as we now have molecules which can hold energy in any amounts, it may be dispensed with, the modification being justified on the same assumptions and by the same sort of limiting process as have been used in this section. Again, we can see that the correct results are obtained if $\bar{H}(z)$ replaces $h(z)$ in (12.2) and all the other integrals, even though the interpretation as coefficient in a power series is no longer possible, and though the integrand is no longer single valued. In such many-valued integrands the limiting process shows that we simply require to take that value which is real on the positive side of the real axis.

§ 13. The Maxwell Distribution Law.

We have carried out the whole process so far with quantized systems included in the assembly, but it may be observed that it is immediately applicable to an assembly composed solely of molecules. If this is done the value of \bar{c}_i in (12.51) establishes at once the Maxwell distribution law, and its fluctuation in (12.55) proves that it is a normal property of the assembly. This is probably the simplest complete proof of the ordinary distribution law; its special advantage is that by means of the fluctuations it is easily established that the actual distribution will hardly ever be far from the average.

The method can also be made to establish the distribution law for a mixture of gases*, and indeed for a mixture of any kind, provided that the systems can be considered to have separate energies.

It is also possible to extend the method to cases in which the total momentum or angular momentum is conserved, by constructing partition functions in more than one independent variable. In fact, there will be as many independent variables as there are uniform integrals of the dynamical equations of the assembly. For simplicity we shall suppose that the linear momentum in a given direction is conserved,

* The effects of the semi-permeable membranes of thermodynamics can be conveniently treated by the partition function.

and let its total amount be G . The method now requires the averaging process to be applied to expressions depending on

$$\frac{P!}{c_1! c_2! \dots} \delta_1^{c_1} \delta_2^{c_2},$$

where we now have not only

$$\sum_t c_t = P, \quad \sum_t c_t \zeta_t = E,$$

but also

$$\sum_t c_t \mu_t = G,$$

where μ_t is the momentum in the given direction of a molecule in the t th cell. To sum the appropriate expressions we must take as our partition function

$$h(z, x) = \sum_t \delta_t z^{\zeta_t} x^{\mu_t}.$$

With this function C will be the coefficient of $z^E x^G$ in $[h(z, x)]^P$, and this can be expressed as a double contour integral. So can the other averages, and the usual asymptotic expansions can be found. The correct distribution law follows on replacing $h(z, x)$ by the integral which is its limit when the sizes of the cells tend to zero. This subject lies rather outside the theme of the present paper and need not be elaborated further.

§ 16. Summary.

The whole paper is concerned with a method of calculating partitions of energy by replacing the usual calculation, which obtains the *most probable* state, and is mathematically unsatisfactory, by a calculation of the *average* state, which is the quantity that is actually required and which can be found with rigour by the use of the multinomial theorem together with a certain theorem in complex variable theory.

After a review of principles and two preliminary examples the real point of the method is illustrated in § 5. Here there are two groups of interacting Planck vibrators of different types. It is shown that the partition can be found by evaluating the coefficient of a certain power of z in an expression which is the product of power series in z . This coefficient can be expressed as a contour integral and can be evaluated by a well-known method, the "method of steepest descents." The result expresses itself naturally in terms of a parameter \mathfrak{S} which is identified with temperature measured on a scale given by $\mathfrak{S} = e^{-1/kT}$.

The work is extended to cover the partition among more

general quantized systems in §9, and examples are given. In §12 it is shown how it may be made to deal with assemblies composed partly of free molecules and partly of quantized systems. In §13 we deal with extensions possible when only molecules are present.

The methods we have described can also be made to throw an interesting light on the statistical foundations of thermodynamics; but in that connexion many points have arisen which require rather careful discussion, and in order not to make the present paper too long, we have deferred them to a future communication.

Cambridge,
May, 1922.

XLV. *The Heterodyne Beat Method and some Applications to Physical Measurements.* By MAURICE H. BELZ, M.Sc. (Cantab.); Barker Graduate Scholar of the University of Sydney*.

IN a recent paper †, a preliminary account was given of the application of the heterodyne beat method to the measurement of magnetic susceptibilities. In virtue of the importance of the method as a sensitive measure of physical quantities, it seems desirable to give a more complete account of the principle and of some of the difficulties encountered in its application.

Essentially the method consists of the following arrangement shown in fig. 1.

Two oscillating circuits, Set 1 and Set 2, are set up side by side and arranged so as to have approximately the same frequency. The two sets are loosely coupled so that in the telephone included in one of the circuits a resultant beat frequency is maintained equal to the difference between the frequencies of the fundamentals or overtones in the two circuits. If symmetry in the two circuits is essential, direct coupling can be replaced by indirect coupling by means of a third circuit in which the telephone is placed. In either case, when the beat frequency is low enough, an audible note will be heard in the telephone, and any changes in the constants of either circuit will cause the frequency of the audible note to alter by an amount equal to the change in frequency of the responsible circuit. This at once provides

* Communicated by Professor Sir E. Rutherford, F.R.S.

† Belz, Proc. Camb. Phil. Soc. vol. xxi. part 2 (1922).