CHAPTER 2

DIGRESSION ON STATISTICAL THERMODYNAMICS

§2.01 *Microdescriptions and macrodescriptions of a system*

According to quantum theory the state of a system is completely specified by its eigenfunction. To each state there corresponds one eigenfunction and to each eigenfunction one state. Such a description of the system we shall call a *microdescription.*

It is often, though not always, possible to regard the system as consisting of a large number of almost independent *units* (molecules, atoms, ions, electrons) and to express each eigenfunction of the system as a linear combination of products of the eigenfunctions of all the units. According to the *symmetry restrictions*, if any, imposed on the eigenfunctions of the system, we then obtain three alternative sets of statistical formulae referred to by the names of Fermi–Dirac, Bose–Einstein, and Boltzmann, respectively. These three alternatives, however, arise only when we express the eigenfunctions of the system in terms of those of the constituent units. As long as we refer only to the eigenfunctions of the whole system, we shall not need to consider these three alternatives separately. Nor shall we do so until we reach §2.10.

When we describe the equilibrium properties of a system by thermodynamic methods, we are not interested in such a precise description as the *microdescription*, but are content with a more crude large scale description, which we shall call a *macrodescription.* For example a possible *macrodescription* of the system would be a precise statement of the energy, the volume, the chemical composition (and in special cases other quantities all measurable on a large scale) of each homogeneous part or phase. For brevity we shall confine our discussion initially to systems whose macrodescription requires a precise statement of only four quantities. The extension of the argument to more complicated systems should be obvious. Initially we shall take the first of these quantities to be the energy, the second to be the volume, the third to be the empirical composition; the nature of the fourth quantity is best indicated by some specific examples.

Example 1 Let us consider a definite quantity of hydrogen (free from deuterium) of given energy and given volume. Then we can complete the description by a statement of what fraction of it is *para*, the remaining fraction being *ortho*.

Example 2 If instead of hydrogen, we have lactic acid we can complete the description by a statement of what fraction is *dextro*, the remaining fraction being *laevo*.

Example 3 If the system consists of a given quantity of iodine of given energy and volume we can complete the description by a statement of what fraction is in the *diatomic* form I_2, the remainder being in the *monatomic* form I.

Example 4 If the system consists of a given quantity of tin of given energy and volume, we can complete the description by stating what fraction is *white*, the remainder being *grey*.

Example 5 If the system consists of a given quantity of sulphur dioxide, we can complete the description by stating what fraction is *liquid*, the remainder being *vapour*.

In the first three examples it is assumed either that the system is homogeneous or, if it consists of two phases, that we are not interested in the relative amounts, these being determined by the other conditions. Another example that might be suggested is a system of a given quantity of hydrogen of given energy and volume for which we were interested both in the ratio of *para* to *ortho* and in the ratio of *liquid* to *vapour*. Such a system, however, requires five quantities, instead of four, to complete its macrodescription and so lies outside the class which we shall discuss, although the extension of the treatment to such a system in fact offers no difficulty.

Having made clear by these examples the nature of the fourth independent variable describing the system we shall denote this variable by ξ. It corresponds closely to the quantity ξ which, following De Donder, we introduced in § 1.44 and which we call the *extent of reaction* of a physico-chemical change. It is not a necessary property of ξ that one should be able completely to control its value, provided that its value can in principle be measured by *macroexperiments.*

§2.02 *System of given energy, volume, and composition*

Let us now consider in more detail a system of prescribed energy U, prescribed volume V, and containing a prescribed number N of molecules of a given kind. Let the number of independent eigenfunctions of the system consistent with the prescribed values of U, V, N and corresponding to a particular value of the parameter ξ be denoted by $\Omega(\xi)$. As long as we are not interested in distinguishing between the states of equal ξ, we may conveniently group them together.

Then the *fundamental assumption of statistical thermodynamics* is the following.

The average properties of the system for prescribed values of U, V, N can be derived statistically by averaging over all groups of states of given ξ, assigning to each group a weight $\Omega(\xi)$.

In other words it is assumed that for given U, V, N the *probability* of a particular value of ξ is

$$\Omega(\xi)/\sum_{\xi}\Omega(\xi). \qquad 2.02.1$$

It is customary to refer to the numerator $\Omega(\xi)$ in (1) as the *thermodynamic probability* of the particular value of ξ. It must be emphasized that *thermodynamic probability* thus defined is not a *probability* in the usual sense of the word. Whereas an ordinary *probability* such as (1) is a number less than or equal to unity, the *thermodynamic probability* is generally a large number.

For reasons which will appear later $\Omega(\xi)$ had better be called the *thermodynamic probability of ξ for given U, V, N* than merely the thermodynamic probability of ξ. Another name for $\Omega(\xi)$ is the *partition function for given U, V, N, ξ*. The reason for this name will also become clearer as we proceed.

We now define a quantity $S(U, V, N, \xi)$ by the relation

$$S(U, V, N, \xi)=k \ln \Omega(\xi) \qquad 2.02.2$$

where k is a universal arbitrary constant whose value will be settled later. It can then be shown as we shall see later that, in a macroscopic system, S has all the properties of the entropy of the system in the macrostate defined by U, V, N, ξ. Formula (2) is a precise formulation of the well-known relation due to Boltzmann to whom the name *thermodynamic probability* is due.

We shall see that Boltzmann's relation (2) between the *entropy* and the *thermodynamic probability* or *partition function* for given U, V, N, ξ is merely one of a number of relations of a similar type between a *characteristic function* for a particular set of variables on the one hand and the *thermodynamic probability* or *partition function* for the same set of variables on the other.

§2.03 *Characteristic of macroscopic system*

According to the *fundamental assumption of statistical thermodynamics* in a system of given U, V, N the average value $\langle\xi\rangle$ of ξ is determined by

$$\langle\xi\rangle=\sum_{\xi}\xi\Omega(\xi)/\sum_{\xi}\Omega(\xi) \qquad 2.03.1$$

and the average value $\langle\xi^2\rangle$ of ξ^2 by

$$\langle\xi^2\rangle=\sum_{\xi}\xi^2\Omega(\xi)/\sum_{\xi}\Omega(\xi). \qquad 2.03.2$$

Thus in general $\langle\xi\rangle^2$ is not the same as $\langle\xi^2\rangle$.

In other words there are fluctuations measured by

$$\begin{aligned}\langle(\xi-\langle\xi\rangle)^2\rangle&=\langle\xi^2\rangle-\langle\xi\rangle^2\\&=\{\sum_{\xi}\Omega(\xi)\sum_{\xi}\xi^2\Omega(\xi)-\sum_{\xi}\xi\Omega(\xi)\sum_{\xi}\xi\Omega(\xi)\}/\{\sum_{\xi}\Omega(\xi)\}^2.\end{aligned} \qquad 2.03.3$$

It can be shown generally that the larger the system the less important is this fluctuation and that for any macroscopic system the fluctuation is entirely trivial compared with $\langle\xi\rangle^2$ itself. Without attempting a proof we shall consider a little more closely how this comes about.

There is some value ξ_{m} of ξ for which $\Omega(\xi)$ has a maximum. Generally speaking the larger the system the sharper is this maximum and for any macroscopic system it is very sharp indeed. On each side of this maximum term $\Omega(\xi_{\text{m}})$ there will be many terms almost as great as $\Omega(\xi_{\text{m}})$. Then there will be a still greater number of terms appreciably smaller but not negligible; but an overwhelming majority of the terms will be entirely negligible, and this majority includes all those terms in which ξ differs appreciably from ξ_{m}.

As a result of such considerations it can be shown that whereas the average properties are strictly determined by attributing to each ξ the weight $\Omega(\xi)$ we may in any macroscopic system with trivial inaccuracy ignore all values of ξ other than the value ξ_{m} at which $\Omega(\xi)$ is maximum.

Thus for any macroscopic system we have with trivial inaccuracy

$$\langle\xi\rangle=\xi_{\text{m}} \qquad 2.03.4$$

$$\langle\xi^2\rangle=\xi_{\text{m}}^2 \qquad 2.03.5$$

and so on.

It is instructive to relate this important characteristic of a macroscopic system to the quantity $S(U, V, N, \xi)$ defined by (2.02.2), namely

$$S(U, V, N, \xi)=k\ln\Omega(\xi). \qquad 2.03.6$$

Let us now define another quantity $S(U, V, N)$ by

$$S(U, V, N)=k\ln\{\sum_{\xi}\Omega(\xi)\}. \qquad 2.03.7$$

Then by definition it is evident that

$$S(U, V, N) > S(U, V, N, \xi) \qquad \text{(all values of } \xi\text{).} \qquad 2.03.8$$

Let us now consider the ratio

$$\{\ln \sum_{\xi} \Omega(\xi) - \ln \Omega(\xi_{\mathrm{m}})\} / \ln \Omega(\xi_{\mathrm{m}}). \qquad 2.03.9$$

It can be shown that roughly speaking $\Omega(\xi_{\mathrm{m}})$ is of the order $N!$ and $\Sigma_{\xi}(\Omega(\xi)/\Omega(\xi_{\mathrm{m}})$ is of the order N^{α} where α is comparable with unity. Hence the numerator in (9) is of the order $\alpha \ln N$ and the denominator of the order $N \ln N$. Thus the expression (9) is of the order α/N or near enough N^{-1}, which is entirely negligible in any macroscopic system. Hence, although the inequality (8) is strictly true by definition for all values of ξ, in any macroscopic system when ξ has the special value ξ_{m} we may with trivial inaccuracy replace the inequality (8) by the equality

$$S(U, V, N) = S(U, V, N, \xi_{\mathrm{m}}). \qquad 2.03.10$$

We shall see in §2.05 that the functions denoted by S have in fact the properties of entropy. Anticipating this identification let us call $S(U, V, N, \xi)$ the *entropy for fixed* ξ and $S(U, V, N)$ the *entropy for equilibrium* ξ.

Consider now a system of given U, V, N with ξ *frozen.* Now suppose that by introduction of a catalyst ξ is *thawed,* so that it takes its equilibrium value. By definition the entropy changes from $S(U, V, N, \xi)$ to $S(U, V, N)$ and also by definition this is always an increase. Only in the special case that the initial value of ξ was ξ_{m} the entropy increase from $S(U, V, N, \xi_{\mathrm{m}})$ to $S(U, V, N)$ for any macroscopic system is trivial. In other words although $S(U, V, N)$ the entropy for *equilibrium* ξ is by definition greater than the entropy for ξ *fixed at its equilibrium value* ξ_{m} the difference in a macroscopic system is negligible and trivial.

We shall see later that a macroscopic system has other characteristics similar and parallel to that just formulated. These characteristics can be summed up in the single sentence that in a *macroscopic* system *fluctuations* of measurable properties are *negligible.*

§2.04 *System of given temperature, volume, and composition*

We shall now consider a system whose volume V and composition N are still prescribed, but instead of prescribing the energy we shall suppose the system to be immersed in a large temperature bath with which it can exchange energy so that the energy of the system can now take various values U_0, U_1, and so on. Let us now enumerate the eigenfunctions of the system for the prescribed values of V and N and for some definite value of ξ;

let there be Ω_r such eigenfunctions corresponding to an energy $U_r(V, N, \xi)$.

From the *fundamental assumption of statistical thermodynamics*, as stated in §2.02, without any further assumptions it can be shown that the average properties of the system in the temperature bath for the prescribed values of V and N can be derived statistically by averaging over all degenerate energy values attaching to each state r of specified ξ and U_r a weight

$$\Omega_r \exp(-\beta U_r) \qquad 2.04.1$$

where β is determined entirely by the temperature bath and so may be regarded as a temperature scale.

The fact that the parameter β is found to appear without any new assumption is the statistical thermodynamic basis of the *zeroth law of classical thermodynamics*. The *statistical thermodynamic* equivalent of the *first law of classical thermodynamics* is merely the principle of *conservation of energy* applied on the microscopic scale, that is to say applied to molecules, atoms, electrons, etc. Thus this principle is from the point of view of statistical thermodynamics not a new law but merely one item in general atomic quantum theory.

To relate the *second law of classical thermodynamics* to *statistical thermodynamics* we make certain algebraic transformations. We begin by defining a function $J(\beta, V, N, \xi)$ by

$$J(\beta, V, N, \xi) = k \ln \{\sum_r \Omega_r(\xi) \exp(-\beta U_r)\} \qquad 2.04.2$$

where the summation is over all states of given ξ, and k is a universal arbitrary constant.

In the system with temperature specified by β there will be fluctuations of U, but the experimentally measurable U will be $\langle U \rangle$, the average value of U. Let us now consider the value of $\langle U \rangle$ for specified β, V, N, and ξ. Using the weighting factors (1) we have

$$\begin{aligned} U(\beta, V, N, \xi) &= \sum_r U_r \Omega_r(\xi) \exp(-\beta U_r) / \sum_r \Omega_r(\xi) \exp(-\beta U_r) \\ &= -\partial \ln\{\sum_r \Omega_r(\xi) \exp(-\beta U_r)\}/\partial\beta \\ &= k^{-1} \partial J(\beta, V, N, \xi)/\partial\beta \qquad 2.04.3 \end{aligned}$$

using (2).

Again associated with the fluctuations in U there will be fluctuations in the pressure $(-\partial U/\partial V)$ but the experimentally measured pressure P will be $\langle -\partial U/\partial V \rangle$. We accordingly have for given β, V, N, and ξ

$$P=\langle -\partial U/\partial V\rangle = -\sum_r (\partial U_r/\partial V)\Omega_r(\xi)\exp(-\beta U_r)/\sum_r \Omega_r(\xi)\exp(-\beta U_r)$$
$$=\beta^{-1}\partial \ln\{\sum_r \Omega_r(\xi)\exp(-\beta U_r)\}/\partial V$$
$$=-k^{-1}\beta^{-1}\partial J(\beta, V, N, \xi)/\partial\beta \qquad 2.04.4$$

using (2).

Let us now make the further algebraic substitution

$$T=k^{-1}\beta^{-1} \qquad 2.04.5$$

and use T as an independent variable instead of β. We now have in place of (3) and (4)

$$\langle U(T, V, N, \xi)\rangle = -\partial J(T, V, N, \xi)/\partial T^{-1}$$
$$=T^2\partial J(T, V, N, \xi)/\partial T \qquad 2.04.6$$

$$P=T\partial J(T, V, N, \xi)/\partial V. \qquad 2.04.7$$

Combining (6) with (7) we have

$$\mathrm{d}J=(\langle U\rangle/T^2)\mathrm{d}T+(P/T)\mathrm{d}V. \qquad 2.04.8$$

Comparing (8) with (1.28.15) we see that the dependence of J, defined by (2), on T defined by (5) and on V is precisely the same as the dependence of the *Massieu function* on the thermodynamic temperature and on the volume. It can in fact be shown that T defined by (5) has all the properties of absolute temperature and J defined by (2) has all the properties of the *Massieu function*. This constitutes a brief summary of how the *second law of classical thermodynamics* follows as a natural deduction from *statistical thermodynamics*.

For the benefit of the reader not familiar with the *Massieu function* J we recall that it is defined by

$$J=S-U/T \qquad 2.04.9$$

and that either the Massieu function J or the Helmholtz function $\mathcal{F}=-TJ$ is a characteristic function for the independent variables T, V, N.

We can now substitute from (5) into (1) and so have as a weighting factor for each energy U_r

$$\Omega_r \exp(-U_r/kT) \qquad 2.04.10$$

and this factor is called the *Boltzmann factor*. From (10) it is clear that kT has the dimensions of energy. k is a universal constant called the *Boltzmann constant*. If we use the *Kelvin scale* of thermodynamic temperature then

$$k=1.3805\times 10^{-23}\,\mathrm{J\ K^{-1}}. \qquad 2.04.11$$

From (10) we see that the average properties of the system for prescribed values of T, V, N and unspecified ξ can be obtained by averaging over all ξ attaching to each ξ a weight $Q(T, \xi)$ defined by

$$Q(T, \xi) = \sum_r \Omega_r(\xi) \exp(-U_r/kT). \qquad 2.04.12$$

The function $Q(T, \xi)$ is usually called the *partition function*, but a more precise name is the *partition function for given T, V, N, ξ*. An alternative name is the *thermodynamic probability for given T, V, N, ξ*.

Substituting from (5) and (12) into (2) we obtain

$$J(T, V, N, \xi) = k \ln Q(T, \xi). \qquad 2.04.13$$

We observe that this relation between the *characteristic function J* and the statistical probability $Q(T, \xi)$ for given T, V, N, ξ is completely analogous to *Boltzmann's relation* (2.02.2) between the characteristic function S and the thermodynamic probability $\Omega(\xi)$ for given U, V, N, ξ.

§2.05 *Further characteristics of macroscopic system*

Let us consider the individual terms of $Q(T, \xi)$ defined by (2.04.12). Let us denote the maximum term by

$$\Omega_{\mathrm{m}} \exp(-U_{\mathrm{m}}/kT) \qquad 2.05.1$$

noting that this Ω_{m} is not the same as the $\Omega(\xi_{\mathrm{m}})$ of §2.03. Generally speaking the larger the system the sharper this maximum and for any macroscopic system it is so sharp that all terms in $Q(T, \xi)$ in which U_r differs appreciably from U_{m} are entirely trivial. Moreover, although the actual number of terms $Q(T, \xi)$ comparable with (1) may be great, the ratio

$$\frac{\ln Q(T, \xi) - \ln\{\Omega_{\mathrm{m}} \exp(-U_{\mathrm{m}}/kT)\}}{\ln\{\Omega_{\mathrm{m}} \exp(-U_{\mathrm{m}}/kT)\}} \qquad 2.05.2$$

is roughly of the order α/N where α is far nearer to unity than to N. Hence in any macroscopic system the ratio (2) is effectively zero and we may therefore replace the definition (2.04.2) of J by

$$J = k \ln\{\Omega_{\mathrm{m}} \exp(-U_{\mathrm{m}}/kT)\}. \qquad 2.05.3$$

It follows again that with an inaccuracy trivial for a macroscopic system

$$\langle U \rangle = T^2 \partial J/\partial T = U_{\mathrm{m}}. \qquad 2.05.4$$

From the classical definition (1.28.13) of the *Massieu function J*, we have

$$S = J + U/T. \tag{2.05.5}$$

We accordingly in *statistical thermodynamics* define a function $S(T, V, N, \xi)$ by

$$S(T, V, N, \xi) = J(T, V, N, \xi) + \langle U \rangle / T. \tag{2.05.6}$$

Using (3), (4), and (5) we obtain from (6)

$$S(T, V, N, \xi) = k \ln \Omega_{\mathrm{m}}(\xi). \tag{2.05.7}$$

Now comparing (7) with (2.02.2) we obtain the striking result

$$S(T, V, N, \xi) = S(\langle U \rangle, V, N, \xi). \tag{2.05.8}$$

Thus although the definition of entropy at a specified temperature by means of (6) together with (2.04.13) is entirely different from the definition of entropy at a specified energy by means of (2.02.2), yet for a macroscopic system the difference between the two is trivial.

This characteristic property of a macroscopic system may be described in the following instructive but less exact way. If we define S by

$$S = k \ln \Omega(\xi) \tag{2.05.9}$$

then in a system of specified energy Ω must denote the number of states having *precisely* this energy, whereas in a system of specified temperature Ω denotes the number of states of energy *nearly* equal to the average energy. The question immediately arises how nearly. The answer is that for a *macroscopic* system it just does not matter.

§2.06 *System of given temperature, pressure, and composition*

We now consider a system of prescribed composition surrounded by a temperature bath and enclosed by a piston subjected to a prescribed pressure P. We construct the double sum

$$W(T, P, N, \xi) = \sum_r \sum_s \Omega_{rs} \exp(-U_r/kT) \exp(-PV_s/kT) \tag{2.06.1}$$

where the summation extends over all energies U_r and all volumes V_s consistent with the prescribed value of ξ. It can then be shown without any new assumptions that we can correctly derive the average (equilibrium) properties of the system for the prescribed values of T, P, N, by averaging over all values of ξ attaching to each a weight $W(T, P, N, \xi)$.

We call $W(T, P, N, \xi)$ the *thermodynamic probability for given* T, P, N, ξ or the *partition function for given* T, P, N, ξ. It is related to the *Planck* *function* Y, which is a characteristic function for the independent variables T, P, N, ξ by

$$Y(T, P, N, \xi)=k \ln W(T, P, N, \xi) \qquad 2.06.2$$

analogous to (2.02.2) and (2.04.13).

For the benefit of the reader unfamiliar with the *Planck function* Y we recall its relation to the *Gibbs function* G, namely

$$Y=-G/T. \qquad 2.06.3$$

Provided the system is macroscopic we may again with only trivial inaccuracy replace W by its maximum term, say

$$\Omega_{\mathrm{m}}(\xi) \exp(-U_{\mathrm{m}}/kT) \exp(-PV_{\mathrm{m}}/kT) \qquad 2.06.4$$

so that we may replace (2) by

$$Y(T, P, N, \xi)=-k \ln \Omega_{\mathrm{m}}(\xi)-U_{\mathrm{m}}/T-PV_{\mathrm{m}}/T. \qquad 2.06.5$$

From (1) and (5) we immediately verify that

$$\langle U+PV\rangle=T^2 \partial Y/\partial T=U_{\mathrm{m}}+PV_{\mathrm{m}} \qquad 2.06.6$$

$$\langle V\rangle=-T\partial Y/\partial P=V_{\mathrm{m}} \qquad 2.06.7$$

as we should expect according to (1.28.16). Furthermore comparing (5) with (1.28.14) we obtain

$$S(T, P, N, \xi)=k \ln \Omega_{\mathrm{m}}(\xi) \qquad 2.06.8$$

verifying that for a macroscopic system the entropy at given T, P is indistinguishable from the entropy at given $U=U_{\mathrm{m}}$ and $V=V_{\mathrm{m}}$.

§2.07 *System of given temperature, pressure, and chemical potential*

To conclude we choose as independent variables the temperature T, pressure P, and chemical potential μ. An illustrative example is a gas in contact with a crystal of the same substance; the crystal is not considered as part of the system. Such a system is called *open*.

We now construct the triple sum

$$W(T, P, \mu, \xi)=\sum_r \sum_s \sum_t \Omega_{rst} \exp(-U_r/kT) \exp(-PV_s/kT) \exp(\mu N_t/LkT) \qquad 2.07.1$$

where L is a general constant called the Avogadro constant defined later in §3.13 and where Ω_{rst} denotes the number of states of energy U_r, volume V_s, and content N_t corresponding to the given value of ξ and the triple summation extends over all sets of values of U_r, V_s, N_t corresponding to the given value of

ξ. It can then be shown without any new assumptions that all the average properties of the system for the prescribed values of T, P, μ are correctly obtained by averaging over all values of ξ attaching to each a weight $W(T, P, \mu, \xi)$ this expression being the *partition function* or *thermodynamic probability of* ξ *for given* T, P, μ.

For a macroscopic system W can in the usual way be replaced by its maximum term say

$$\Omega_{\mathrm{m}}(\xi) \exp(-U_{\mathrm{m}}/kT) \exp(-PV_{\mathrm{m}}/kT) \exp(\mu N_{\mathrm{m}}/LkT). \qquad 2.07.2$$

If we now define a quantity $O(T, P, \mu, \xi)$ by

$$O(T, P, \mu, \xi) = k \ln W(T, P, \mu, \xi) \qquad 2.07.3$$

we may for a macroscopic system replace (3) by

$$O(T, P, \mu, \xi) = k \ln \Omega_{\mathrm{m}}(\xi) - U_{\mathrm{m}}/T - PV_{\mathrm{m}}/T + \mu N_{\mathrm{m}}/LT. \qquad 2.07.4$$

Moreover for a macroscopic system we have as usual

$$S = k \ln \Omega_{\mathrm{m}}(\xi) \qquad 2.07.5$$

$$\langle U \rangle = U_{\mathrm{m}} \qquad 2.07.6$$

$$\langle V \rangle = V_{\mathrm{m}} \qquad 2.07.7$$

$$\langle N \rangle = N_{\mathrm{m}}. \qquad 2.07.8$$

Comparing (5) to (8) with (4), dropping subscripts and replacing N/L by n we find that

$$O(T, P, N, \xi) = S - U/T - PV/T + \mu n/T = 0 \qquad 2.07.9$$

according to (1.28.5).

From the analogy between (3), (2.02.2), (2.04.13) and (2.06.2) we expect $O(T, P, \mu, \xi)$ to be a *characteristic function* for the variables T, P, and μ. According to (9) this characteristic function is identically zero. We now recall the Gibbs–Duhem relation (1.30.2)

$$S\mathrm{d}T - V\mathrm{d}P + \sum_i n_i \mathrm{d}\mu_i = 0. \qquad 2.07.10$$

In a system of one component the sum $\Sigma_i\, n_i \mathrm{d}\mu_i$ reduces to $n\mathrm{d}\mu$ and so (10) becomes

$$0 = S\mathrm{d}T - V\mathrm{d}P + n\mathrm{d}\mu \qquad 2.07.11$$

showing that the characteristic function for the independent variables T, P, μ is indeed zero.

§2.08 *Recapitulation*

We can now summarize the content of the several preceding sections*. For each selected set of three independent variables, other than ξ, a different kind of weighting factor w has to be attached to the microstates. The sum Σw for all microstates consistent with the prescribed values of the three chosen independent variables other than ξ and corresponding to a definite value of ξ is called the *partition function* or the *thermodynamic probability* for the prescribed values of ξ and the other three independent variables. Furthermore in each case $k \ln(\Sigma w)$ is a *characteristic function* for the chosen set of three independent variables other than ξ. These relationships are shown in table 2.1.

TABLE 2.1

Independent variables	Weighting factor for each microstate	Characteristic function equal to $k \ln(\Sigma w)$
U, V, N, ξ	1	S
T, V, N, ξ	$\exp(-U/kT)$	$J = -\mathcal{F}/T$
T, P, N, ξ	$\exp(-U/kT)\exp(-PV/kT)$	$Y = -G/T$
T, P, μ, ξ	$\exp(-U/kT)\exp(-PV/kT)\exp(\mu N/LkT)$	zero

We emphasize again that each of the listed *characteristic functions* S, J, Y, and zero is related to the corresponding *thermodynamic probability* according to

$$\text{characteristic function} = k \ln(\text{thermodynamic probability}).$$

The earliest and best known example of this form is Boltzmann's relation for $S(U, V, N, \xi)$, but other examples and particularly that for $J(T, V, N, \xi)$ are in fact more useful.

It is a fundamental characteristic of a macroscopic system that any partition function may with trivial inaccuracy be replaced by its maximum term. It follows that the equilibrium value of ξ is that value which maximizes the characteristic function belonging to the chosen set of independent variables. The alternative equilibrium conditions

for given U and V that S is a *maximum* 2.08.1

for given T and V that J is a *maximum* 2.08.2

for given T and P that Y is a *maximum* 2.08.3

---

* Guggenheim, J. Chem. Phys. 1939 7 103; Forh. 5te Nordiske Kemikermøde København 1939 p. 205.

thus obtained are precisely equivalent to (1.35.1), (1.35.3), and (1.35.4) respectively.

§2.09 *Extension to several components. Absolute activities*

We have hitherto restricted our exposition to systems of a single component purely for the sake of brevity. The extension to systems of several components is straightforward.

In particular for a system at given values of the independent variables T, P, and the μ_i's the weighting factor for each independent microstate will be

$$\exp(-U/kT)\exp(-PV/kT)\prod_i \lambda_i^{N_i} \qquad 2.09.1$$

where for brevity we have introduced quantities λ_i defined by

$$\lambda_i = \exp(\mu_i/LkT)$$

or

$$\mu_i = LkT \ln \lambda_i . \qquad 2.09.2$$

These quantities λ_i may be used instead of the μ_i and are often more convenient. λ_i is called the *absolute activity* of the species *i*. We shall meet these quantities again in §3.15.

§2.10 *Antisymmetric and symmetric eigenfunctions*

In §2.01 we mentioned that it is often, though not always, possible to regard the units (molecules, atoms, ions, electrons) composing the system as almost independent. In this case each eigenfunction of the system can be expressed as a linear combination of products of the eigenfunctions of all the units. We begin by considering the case that all the units are of the same kind. We denote the eigenfunctions of the units by ϕ and the eigenfunctions of the whole system by ψ. We have now to distinguish two cases.

If each unit is a fundamental particle (proton, neutron, or electron) or is composed of an odd number of fundamental particles, then each eigenfunction ψ of the system is constructed by forming a determinant of the eigenfunctions of the individual units. For the sake of simplicity and brevity we consider a system consisting of only three units, numbered 1, 2, 3. The symbol $\phi_\alpha(1)$ then denotes the eigenfunction of the unit 1 when in the state α. The eigenfunction is then constructed as follows

$$\psi_{\alpha\beta\gamma} = \begin{vmatrix} \phi_\alpha(1) & \phi_\beta(1) & \phi_\gamma(1) \\ \phi_\alpha(2) & \phi_\beta(2) & \phi_\gamma(2) \\ \phi_\alpha(3) & \phi_\beta(3) & \phi_\gamma(3) \end{vmatrix} . \qquad 2.10.1$$

We notice that if we interchange the states of any two units, ψ changes sign. We accordingly describe the eigenfunctions ψ as *antisymmetric* with respect to every pair of units. It follows at once that if any two of the states, α, β, γ are identical then $\psi_{\alpha\beta\gamma}$ vanishes. Thus there is one independent ψ for each combination of three ϕ_α, ϕ_β, ϕ_γ provided α, β, γ are all different but none if any two of α, β, γ are the same.

If on the other hand each unit is a photon or is composed of an even number of fundamental particles (protons, neutrons, electrons), then each eigenfunction of the system is constructed from the eigenfunctions of the units by forming linear combinations called *permanents* similar to determinants, but in which all the terms are added. Thus in the case of only three units 1, 2, 3 the eigenfunction $\psi_{\alpha\beta\gamma}$ is defined by

$$\psi_{\alpha\beta\gamma} = \left\Vert \begin{matrix} \phi_\alpha(1) & \phi_\beta(1) & \phi_\gamma(1) \\ \phi_\alpha(2) & \phi_\beta(2) & \phi_\gamma(2) \\ \phi_\alpha(3) & \phi_\beta(3) & \phi_\gamma(3) \end{matrix} \right\Vert \qquad 2.10.2$$

which differs from (1) in that all the six terms are added. We notice that if we interchange the states of any two units, ψ remains unchanged. We accordingly describe the eigenfunction ψ as *symmetric* in all the units. It is clear that there is one independent ψ for every combination of three eigenfunctions ϕ_α, ϕ_β, ϕ_γ whether or not any two or more of α, β, γ are the same.

§2.11 *Fermi–Dirac and Bose–Einstein statistics*

Let us now consider a system containing N indistinguishable units and enquire how many eigenfunctions ψ of the system can be constructed out of g eigenfunctions ϕ of the units. There are two distinct problems with different answers according as ψ is to be antisymmetric or symmetric in the units.

In the case where ψ is to be antisymmetric, to obtain any such ψ at all, g must be at least as great as N and the number of such eigenfunctions ψ is then

$$g!/N!(g-N)! \qquad \text{(antisymmetric, } g \geqq N\text{)}. \qquad 2.11.1$$

In the other case where ψ is to be symmetric, the number of such eigenfunctions ψ is

$$(g+N-1)!/(g-1)!N! \qquad \text{(symmetric)} \qquad 2.11.2$$

which, when $g \gg 1$, differs only trivially from the simpler expression

$$(g+N)!/g!N!. \qquad 2.11.3$$

It is of interest to note that when $g \gg N$, both (1) and (3) are nearly the same as

$$g^N/N! \qquad (g \gg N). \tag{2.11.4}$$

If now we translate the laws governing the average properties of the whole system outlined in §§2.01–2.09 into forms relating to the average distributions of the component units, we shall as a consequence of the difference between (1) and (2) find different results according as the eigenfunctions ψ are to be antisymmetric or symmetric in the units. These distribution laws take the simplest form if we choose as independent variables the temperature T, the volume V, and the absolute activity λ. We shall now state these laws without derivation.

Let ε_α denote the energy of a unit in the state α having the eigenfunction ϕ_α. Then if the unit is a fundamental particle (proton, neutron, or electron) or is composed of an odd number of fundamental particles, the eigenfunction ψ must be antisymmetric in the units and the average number N_α of units in the state α is found to be given by

$$N_\alpha/(1-N_\alpha)=\lambda \exp(-\varepsilon_\alpha/kT) \tag{2.11.5}$$

where λ denotes the *absolute activity* of the unit, T the absolute temperature, and k the Boltzmann constant. This distribution law is called that of *Fermi–Dirac statistics.*

If on the other hand the unit is a photon or is composed of an even number of fundamental particles, the eigenfunction ψ must be symmetric in the units and the average number N_α of units in the state α is found to be given by

$$N_\alpha/(1+N_\alpha)=\lambda \exp(-\varepsilon_\alpha/kT). \tag{2.11.6}$$

This distribution law is called that of *Bose–Einstein statistics.*

It is to be noted that in both the cases of *Fermi–Dirac statistics* and *Bose–Einstein statistics* the average number N_α of units in each state is related simply and explicitly to the temperature T and the absolute activity λ, which we recall is related to the chemical potential μ by (2.09.2).

§2.12 *Boltzmann statistics*

Let the subscript $_0$ denote the state of lowest energy ε_0 and let us consider the case that

$$\lambda \exp(-\varepsilon_0/kT) \ll 1 \tag{2.12.1}$$

so that a fortiori

$$\lambda \exp(-\varepsilon_\alpha/kT) \ll 1 \qquad (\text{all } \alpha). \tag{2.12.2}$$

It then follows from either (2.11.5) or (2.11.6) that

$$N_\alpha \ll 1 \qquad (\text{all } \alpha). \tag{2.12.3}$$

We may then without loss of accuracy replace either (2.11.5) or (2.11.6) by

$$N_\alpha = \lambda \exp(-\varepsilon_\alpha/kT). \tag{2.12.4}$$

This distribution law is called that of *Boltzmann statistics.*

We now state without proof that in almost all the systems met in practice the condition (1) is satisfied. There are only two important exceptions. The first is the system of conducting electrons in a metal; these obey the Fermi–Dirac distribution law and will not be discussed in this book. The other is the system of photons forming radiation; these obey the Bose–Einstein distribution law and will be discussed in chapter 12. Boltzmann statistics are sufficient for all the other systems to be met in this book and from here onwards we shall confine our attention to these.

§2.13 *Partition functions of units and thermodynamic functions*

For any system obeying Boltzmann statistics, we have according to (2.12.4)

$$N_\alpha = \lambda \exp(-\varepsilon_\alpha/kT). \tag{2.13.1}$$

If we apply (1) to every state and add, we obtain

$$N = \lambda \sum_\alpha \exp(-\varepsilon_\alpha/kT) \tag{2.13.2}$$

so that

$$\mu/LkT = \ln \lambda = \ln\{N/\sum_\alpha \exp(-\varepsilon_\alpha/kT)\}. \tag{2.13.3}$$

The sum $\Sigma_\alpha \exp(-\varepsilon_\alpha/kT)$ is called the *partition function of the units.* Its structure is similar to that of the *partition function of the whole system* for the independent variables T, V, N. Formula (3) is the basis for the evaluation of the thermodynamic functions in terms of the energies of all the states of the component units.

Formula (3) is equivalent to the formula for the *Massieu function J*

$$J = -\mathit{Æ}/T = S - U/T = k \ln[\{\sum_\alpha \exp(-\varepsilon_\alpha/kT)\}^N/N!]. \tag{2.13.4}$$

If we compare (4) with (2.04.13) we see that the two are equivalent when we bear in mind that the factor $N!$ in the denominator in (4) is required to avoid counting as distinct states those obtainable from one another by a mere permutation of indistinguishable units.

The more general formula for a system containing more than one kind of units (molecules) is

$$J = -\mathit{F}/T = S - U/T = k\sum_i \ln[\{\sum_\alpha \exp(-\varepsilon_\alpha/kT)\}^{N_i}/N_i!]. \qquad 2.13.5$$

§2.14 *Separable degrees of freedom*

It is often the case that there is no appreciable interaction between two or more degrees of freedom of a unit. Such degrees of freedom are said to be *separable*. Each eigenfunction ϕ_α may then be expressed as a product of the eigenfunctions for the several separable degrees of freedom, and the energy ε_α as the sum of the energies of the several separable degrees of freedom. It then follows immediately that the partition function of the unit can be expressed as the product of partition functions for its several separable degrees of freedom.

In particular the translational degrees of freedom of molecules are usually separable from the internal degrees of freedom. Among the internal degrees of freedom we here include rotational degrees of freedom as well as atomic vibrations and electronic and nuclear degrees of freedom. We may accordingly write for the partition function of a molecule

$$\sum \exp(-\varepsilon_\alpha/kT) = \sum \exp(-\varepsilon_{\mathrm{tr}}/kT) \sum \exp(-\varepsilon_{\mathrm{int}}/kT) \qquad 2.14.1$$

where $\varepsilon_{\mathrm{tr}}$ denotes the energy of the translational degrees of freedom and $\varepsilon_{\mathrm{int}}$ the energy of the internal degrees of freedom. Substituting (1) into (2.13.5) we obtain for the Massieu function J and the Helmholtz function F

$$-TJ = \mathit{F} = -kT\sum_i \ln[\{\sum \exp(-\varepsilon_{\mathrm{tr}}/kT) \sum \exp(-\varepsilon_{\mathrm{int}}/kT)\}^{N_i}/N_i!]. \qquad 2.14.2$$

Alternatively we may write

$$J = J_{\mathrm{tr}} + J_{\mathrm{int}} \qquad 2.14.3$$

$$\mathit{F} = \mathit{F}_{\mathrm{tr}} + \mathit{F}_{\mathrm{int}} \qquad 2.14.4$$

$$\mathit{F}_{\mathrm{tr}} = -TJ_{\mathrm{tr}} = -kT\sum_i \ln[\{\sum \exp(-\varepsilon_{\mathrm{tr}}/kT)\}^{N_i}/N_i!] \qquad 2.14.5$$

$$\mathit{F}_{\mathrm{int}} = -TJ_{\mathrm{int}} = -kT\sum_i N_i \ln\{\sum \exp(-\varepsilon_{\mathrm{int}}/kT)\} \qquad 2.14.6$$

where the subscript $_{\mathrm{tr}}$ refers throughout to contributions from the translational degrees of freedom and the subscript $_{\mathrm{int}}$ to contributions from the internal degrees of freedom.

§2.15 *Classical and unexcited degrees of freedom*

It may happen that there are many energy levels less than kT. When this is the case, the sum which defines the partition function may without loss of accuracy be replaced by an integral, whose evaluation is often elementary. Such a degree of freedom is called a *classical degree of freedom.* Whether a particular degree of freedom is classical depends on the temperature. Under ordinary conditions the translational and rotational degrees of freedom of the molecules in a gas are classical.

In the opposite case it may happen that the separation between the states of lowest energy level and those of the next energy level is several times greater than kT. The partition function then reduces effectively to the terms corresponding to the lowest energy level, that is to

$$g_0 \exp(-\varepsilon_0/kT) \tag{2.15.1}$$

where ε_0 denotes the lowest energy level and g_0 denotes the number of states having this energy. Such degrees of freedom are called *unexcited degrees of freedom.* The contribution of each such unexcited degree of freedom to the Helmholtz function $\mathcal{F}$ is clearly

$$\varepsilon_0 - kT \ln g_0 \tag{2.15.2}$$

and the corresponding contribution to the entropy

$$k \ln g_0 \tag{2.15.3}$$

which we notice is independent of the temperature. Whether a particular degree of freedom is unexcited depends by definition on the temperature. At all the temperatures with which we are concerned all degrees of freedom internal to the atomic nucleus are unexcited. The electronic degrees of freedom of most molecules may also be regarded as unexcited at most of the temperatures which concern us; there are however a few exceptions, notably the molecule NO.

§2.16 *Translational degrees of freedom*

The translational degrees of freedom of a dilute gas may be regarded as classical. When the partition function for the translational degrees of freedom of a molecule is replaced by an integral and the integration is performed, one obtains

$$(2\pi mkT/h^2)^{\frac{3}{2}} V \tag{2.16.1}$$

where m denotes the mass of a molecule and V the volume in which it is enclosed; h denotes the Planck constant and k as usual the Boltzmann constant. Thus for a dilute gaseous mixture according to (2.14.5) we have

$$\mathcal{F}_{\mathrm{tr}} = -TJ_{\mathrm{tr}} = -kT \sum_i \ln\{(2\pi mkT/h^2)^{\frac{3}{2}N_i} V^{N_i}/N_i!\} \quad \text{(dilute gas).} \tag{2.16.2}$$

Let us now consider the translational degrees of freedom in a crystal. We may regard each molecule as vibrating about an equilibrium position in the crystal lattice. Let us denote by q the partition function for a molecule attached to a given lattice position and for the moment let us imagine all the N molecules to be individually distinguishable but sufficiently alike so that any one can be interchanged with any other without destroying the crystal structure. Then the molecules can be permuted over the lattice positions in $N!$ ways, so that the partition function for the translational motion of the molecules of the whole crystal would be $N!q^N$. Actually the molecules are of course not individually distinguishable and we must consider only states whose eigenfunction is symmetric in molecules containing an even number of fundamental particles and antisymmetric in molecules containing an odd number of fundamental particles. In the simplest case when all the molecules in the crystal are of the same kind the number of states is thus reduced by a factor $N!$, which cancels the other $N!$, so that the partition function for the whole crystal becomes q^N. We thus have for a crystal of a pure substance

$$\text{\AE}_{\mathrm{tr}} = -TJ_{\mathrm{tr}} = -NkT \ln q \qquad \text{(crystal).} \tag{2.16.3}$$

Each molecule at a given lattice position usually has only one state of lowest translational energy and so at low temperatures q tends to $\exp(-\varepsilon_0/kT)$. We therefore have for a crystal of a pure substance

$$\text{\AE}_{\mathrm{tr}} \to N\varepsilon_0 \qquad (T \to 0) \tag{2.16.4}$$

and consequently

$$S_{\mathrm{tr}} \to 0 \qquad (T \to 0). \tag{2.16.5}$$

For a mixed crystal containing several distinguishable kinds of molecules, e.g. isotopes, the eigenfunctions have to be symmetric, or antisymmetric, only with respect to identical molecules. Hence we have to divide only by the product of all the $N_i!$ instead of by $N!$. We therefore have instead of (3)

$$\text{\AE}_{\mathrm{tr}} = -TJ_{\mathrm{tr}} = -kT \ln N! - kT \sum_i \ln\{q_i^{N_i}/N_i!\} \tag{2.16.6}$$

where $N = \Sigma_i N_i$. It has been implicitly assumed that interchanging two molecules of different kinds in the crystal does not affect the partition function q_i of either of them. This assumption is justified provided the molecules are sufficiently similar, e.g. isotopic. Since at low temperatures each q_i tends to $\exp(-\varepsilon_{i0}/kT)$ it follows that

$$(\text{\AE}_{\mathrm{tr}} - \sum_i N_i\varepsilon_{i0})/kT \to -\ln N! + \sum_i \ln N_i! \qquad (T \to 0) \tag{2.16.7}$$

and consequently

$$S_{\mathrm{tr}} \to k \ln N! - k \sum_i \ln N_i! \qquad (T \to 0). \tag{2.16.8}$$

§2.17 *Third law of thermodynamics*

After this brief and necessarily incomplete sketch of statistical thermodynamics we recall the formulation of the *third law of thermodynamics* which we adopted in §1.66.

By the standard methods of statistical thermodynamics it is possible to derive for certain entropy changes general formulae which cannot be derived from the zeroth, first, or second laws of thermodynamics. In particular we can obtain formulae for entropy changes in highly disperse systems (i.e. dilute gases), those in very cold systems (i.e. when $T \to 0$), and those associated with the mixing of very similar substances (e.g. isotopes).

We shall now briefly state these deductions from statistical thermodynamics without giving detailed derivations.

In the first place we consider the translational term in the thermodynamic functions of a highly disperse system, i.e. a dilute gas, containing N_i molecules of type i having a mass m_i. The contributions to the Helmholtz function $Æ$ and to the Massieu function J are given by

$$Æ_{\mathrm{tr}} = -TJ_{\mathrm{tr}} = -kT \sum_i \ln\{(2\pi m_i kT/h^2)^{\frac{3}{2}N_i} V^{N_i}/N_i!\}. \tag{2.17.1}$$

The corresponding contribution S_{tr} to the entropy S is

$$S_{\mathrm{tr}} = k \sum_i \ln\{(2\pi m_i kT/h^2)^{\frac{3}{2}N_i} V^{N_i}/N_i!\} + k \sum_i \tfrac{3}{2} N_i. \tag{2.17.2}$$

In particular in a gaseous single substance

$$S_{\mathrm{tr}} = k \ln\{(2\pi mkT/h^2)^{\frac{3}{2}N} V^N/N!\} + \tfrac{3}{2}Nk. \tag{2.17.3}$$

Using Stirling's formula for large N

$$\ln N! = N \ln N - N \tag{2.17.4}$$

we can rewrite (3) as

$$S_{\mathrm{tr}}/Nk = \ln\{(2\pi mkT/h^2)^{\frac{3}{2}} V/N\} + \tfrac{5}{2}. \tag{2.17.5}$$

Anticipating the formula given in §§ 3.13–3.14 for the pressure P of a single perfect gas

$$P = NkT/V \tag{2.17.6}$$

we can replace (5) by

$$S_{\mathrm{tr}}/Nk = \ln\{(2\pi m/h^2)^{\frac{3}{2}}(kT)^{\frac{5}{2}}/P\} + \tfrac{5}{2}. \qquad 2.17.7$$

We shall use the equivalent of formula (7) in §3.26.

Our second example is the translational term in the entropy of a crystal of a pure substance. As the temperature tends towards zero, this contribution tends to zero. We shall return to this result in §3.51.

Finally we consider the entropy of mixtures of very similar substances such as isotopes. If several very similar substances, such as isotopes, all at the same temperature and same number of molecules per unit volume are mixed, the temperature and number of molecules per unit volume being kept unchanged, the entropy is increased by ΔS given by

$$\Delta S/k = \ln N! - \sum_i \ln N_i! \qquad 2.17.8$$

where N_i denotes the number of molecules of the species i and $N = \Sigma_i N_i$ denotes the total number of molecules of all species. Using Stirling's formula (4), we can rewrite (8) as

$$\Delta S/k = \sum_i N_i \ln(N/N_i). \qquad 2.17.9$$

This applies to solids, and incidentally to liquids, as well as to gases, provided the various species are sufficiently similar, e.g. isotopic. We shall make use of this in §3.55.

When we meet these formulae again in chapter 3, the number of molecules N_i will be replaced by the amount of substance $n_i = N_i/L$ and correspondingly the Boltzmann constant k will be replaced by the gas constant $R = Lk$.