THE INTERNATIONAL
ENCYCLOPEDIA
OF PHYSICAL CHEMISTRY
AND CHEMICAL PHYSICS

THE INTERNATIONAL ENCYCLOPEDIA OF PHYSICAL CHEMISTRY AND CHEMICAL PHYSICS

Topic 8. STATISTICAL MECHANICS

EDITOR: J. E. MAYER

Volume 1

EQUILIBRIUM STATISTICAL MECHANICS

BY

J. E. MAYER

THE INTERNATIONAL ENCYCLOPEDIA OF PHYSICAL CHEMISTRY AND CHEMICAL PHYSICS

Members of the Honorary Editorial Advisory Board

J. N. Agar, *Cambridge*
R. M. Barrer, *London*
C. E. H. Bawn, *Liverpool*
N. S. Bayliss, *Western Australia*
R. P. Bell, *Oxford*
C. J. F. Böttcher, *Leyden*
F. P. Bowden, *Cambridge*
G. M. Burnett, *Aberdeen*
J. A. V. Butler, *London*
C. A. Coulson, *Oxford*
J. S. Courtney-Pratt, *New Jersey*
D. P. Craig, *London*
F. S. Dainton, *Nottingham*
C. W. Davies, *London*
B. V. Derjaguin, *Moscow*
M. J. S. Dewar, *Texas* (*Austin*)
G. Duyckaerts, *Liège*
D. D. Eley, *Nottingham*
H. Eyring, *Utah*
P. J. Flory, *Mellon Institute*
R. M. Fuoss, *Yale*
P. A. Giguère, *Laval*
W. Groth, *Bonn*
J. Guéron, *Brussels*
C. Kemball, *Queen's, Belfast*
J. A. A. Ketelaar, *Amsterdam*
G. B. Kistiakowsky, *Harvard*
H. C. Longuet-Higgins, *Cambridge*
R. C. Lord, *Massachusetts Institute of Technology*
M. Magat, *Paris*
R. Mecke, *Freiburg*
Sir Harry Melville, *S.R.C.*, *London*
S. Mizushima, *Tokyo*
R. S. Mulliken, *Chicago*
R. G. W. Norrish, *Cambridge*
R. S. Nyholm, *London*
J. T. G. Overbeek, *Utrecht*
K. S. Pitzer, *Rice University, Houston*
J. R. Platt, *Chicago*
G. Porter, *The Royal Institution* (*London*)
I. Prigogine, *Brussels* (*Free University*)
R. E. Richards, *Oxford*
Sir Eric Rideal, *London*
J. Monteath Robertson, *Glasgow*
E. G. Rochow, *Harvard*
G. Scatchard, *Massachusetts Institute of Technology*
Glenn T. Seaborg, *California* (*Berkeley*)
N. Sheppard, *East Anglia*
R. Smoluchowski, *Princeton*
H. Stammreich, *São Paulo*
Sir Hugh Taylor, *Princeton*
H. G. Thode, *McMaster*
H. W. Thompson, *Oxford*
D. Turnbull, *G.E.*, *Schenectady*
H. C. Urey, *California* (*La Jolla*)
E. J. W. Verwey, *Philips, Eindhoven*
B. Vodar, *Laboratoire de Bellvue, France*
M. Kent Wilson, *Tufts*
Lord Wynne-Jones, *Newcastle-upon-Tyne*

THE INTERNATIONAL ENCYCLOPEDIA OF PHYSICAL CHEMISTRY AND CHEMICAL PHYSICS

List of Topics and Editors

1. Mathematical Techniques	H. Jones, *London*
2. Classical and Quantum Mechanics	R. McWeeny, *Keele*
3. Electronic Structure of Atoms	C. A. Hutchison, Jr., *Chicago*
4. Molecular Binding	J. W. Linnett, *Cambridge*
5. Molecular Properties	
(*a*) Electronic	J. W. Linnett, *Cambridge*
(*b*) Non-electronic	N. Sheppard, *East Anglia*
6. Kinetic Theory of Gases	E. A. Guggenheim, *Reading*
7. Classical Thermodynamics	D. H. Everett, *Bristol*
8. Statistical Mechanics	J. E. Mayer, *La Jolla*
9. Transport Phenomena	J. C. McCoubrey, *Birmingham*
10. The Fluid State	J. S. Rowlinson, *London*
11. The Ideal Crystalline State	M. Blackman, *London*
12. Imperfections in Solids	Editor to be appointed
13. Mixtures, Solutions, Chemical and Phase Equilibria	M. L. McGlashan, *Exeter*
14. Properties of Interfaces	D. H. Everett, *Bristol*
15. Equilibrium Properties of Electrolyte Solutions	R. A. Robinson, *Washington, D.C.*
16. Transport Properties of Electrolytes	R. H. Stokes, *Armidale*
17. Macromolecules	C. E. H. Bawn, *Liverpool*
18. Dielectric and Magnetic Properties	J. W. Stout, *Chicago*
19. Gas Kinetics	A. F. Trotman-Dickenson, *Aberystwyth*
20. Solution Kinetics	R. M. Noyes, *Eugene*
21. Solid and Surface Kinetics	F. C. Tompkins, *London*
22. Radiation Chemistry	R. S. Livingston, *Minneapolis*

EQUILIBRIUM STATISTICAL MECHANICS

BY

J. E. MAYER
UNIVERSITY OF CALIFORNIA

PERGAMON PRESS
OXFORD · LONDON · EDINBURGH · NEW YORK
TORONTO · SYDNEY · PARIS · BRAUNSCHWEIG

Pergamon Press Ltd., Headington Hill Hall, Oxford
4 & 5 Fitzroy Square, London W.1
Pergamon Press (Scotland) Ltd., 2 & 3 Teviot Place, Edinburgh 1
Pergamon Press Inc., 44–01 21st Street, Long Island City, New York 11101
Pergamon of Canada Ltd., 207 Queen's Quay West, Toronto 1
Pergamon Press (Aust.) Pty. Ltd., 19a Boundary Street, Rushcutters Bay,
N.S.W. 2011, Australia
Pergamon Press S.A.R.L., 24 rue des Écoles, Paris 5e
Vieweg & Sohn GmbH, Burgplatz 1, Braunschweig

First edition 1968

Library of Congress Catalog Card No. 68–21387

PRINTED IN GREAT BRITAIN AT THE PITMAN PRESS, BATH

08 012647 2

INTRODUCTION

THE International Encyclopedia of Physical Chemistry and Chemical Physics is a comprehensive and modern account of all aspects of the domain of science between chemistry and physics, and is written primarily for the graduate and research worker. The Editors-in-Chief, Professor D. D. ELEY, Professor J. E. MAYER and Professor F. C. TOMPKINS, have grouped the subject matter in some twenty groups (General Topics), each having its own editor. The complete work consists of about one hundred volumes, each volume being restricted to around two hundred pages and having a large measure of independence. Particular importance has been given to the exposition of the fundamental bases of each topic and to the development of the theoretical aspects; experimental details of an essentially practical nature are not emphasized although the theoretical background of techniques and procedures is fully developed.

The Encyclopedia is written throughout in English and the recommendations of the International Union of Pure and Applied Chemistry on notation and cognate matters in physical chemistry are adopted. Abbreviations for names of journals are in accordance with *The World List of Scientific Periodicals.*

CONTENTS

CHAPTER 5 SOLIDS

CHAPTER 6 DENSE FLUIDS

CHAPTER 7 QUANTUM STATISTICS

CHAPTER 8 DERIVATION OF THE PARTITION FUNCTION EQUATIONS

CHAPTER 1

A MASTER EQUATION AND TWO PARTITION FUNCTIONS

1.1. The scope of statistical mechanics

The term statistical mechanical is used to describe the theory by which one attempts to predict and explain the connection between the macroscopic behavior of bulk matter and the microscopic properties of its constituent atoms and molecules. The material with which one deals, usually described by giving the numbers of molecules of each constituent species, and the thermodynamic parameters necessary to give its state, is called the "System." No sharp minimum size is necessarily given, but it is usually tacitly understood that the system consists of something of the order of 10^{19} or more molecules. The macroscopic behavior of the system is measured by quantities like pressure, total volume, temperature, total energy, or any other thermodynamic quantity such as dielectric constant or magnetic susceptibility. Measurements of quantities such as these always are affected by the behavior of at least an appreciable fraction of all the very large number of molecules composing the system. These thermodynamic quantities are the macroscopic variables and properties.

By microscopic properties are meant the mechanical properties of the individual molecules. These include the description of the various stationary quantum states of the molecule, in particular the energies above the lowest energy state, of the rotational, vibrational, and electronic excited states. They also include the law of the forces which act between two or more molecules. The "size" of the molecules is implicitly given by this law of force between molecules. The distance at which the force between two molecules becomes strongly repulsive is approximately the closest distance which they can approach each other in normal bulk matter.

It is not the problem of statistical mechanics to determine these microscopic properties from *a priori* calculation using quantum mechanical principles, but to assume that this problem has already been solved, and to use the result in the prediction of the macroscopic properties. In practice, for very many problems at least, a rather rudimentary knowledge of the properties of the molecules suffices to predict properties with moderate precision. If no excited electronic levels exist with an excitation energy of less than one electron volt, as is most frequently the case for chemically stable molecules, and if the "degeneracy," or number of states in the lowest electronic level is known, this alone is all that is required in knowledge of the electronic structure for most computations below several thousand degrees Kelvin. The rotational levels of diatomic molecules are usually adequately given by a simple equation involving only the moment of inertia of the molecule. For polyatomic molecules a knowledge of the three principal moments of inertia, rather than a knowledge of the rather complicated energy spectrum of the rotational levels, is usually all that is required. The energy spacing of vibrational levels in molecules is given simply in terms of the fundamental frequencies, and these are sought from spectroscopic tabulations rather than from fundamental quantum mechanical calculations. The forces between molecules are actually usually poorly known. Those molecules which have nearly spherical symmetry, including in this category the diatomic elements N_2, O_2, etc., as well as the noble gases and most ions which are truly spherical, are usually assumed to interact with a simple two parameter force dependent only on the distance of their centers. The force between many molecules is assumed to be the sum of the forces between all pairs.

Fortunately for the status of the theory of statistical mechanics most of the formal development can be made without recourse to any doubtful assumptions about the details of the microscopic properties of the molecules. The form of the answers obtained does not usually depend on the exact shape or magnitude of the intermolecular forces, nor on the exact values of the parameters describing the specific molecules. These values enter only in the end determination of the numerical values of the thermodynamic functions. In a very considerable number of cases these characteristics of the molecules are sufficiently accurately known to derive theoretical values for the macroscopic properties of the system which are comparable, or even exceed, in precision the macroscopic measurements which can be made in the laboratory.

The general method of statistical mechanics is such that one should expect it to be equally applicable to systems in thermodynamic equilibrium, and to systems in a state of flux for which the macroscopic properties are changing with time. In actual practice, there is, at least at present, a great difference in the ease of approach, and in the generality with which equations can be written, for the two cases. Statistical mechanics therefore is naturally divided into two fields, equilibrium statistical mechanics, to which the word statistical mechanics is frequently considered to be limited, and the statistical mechanics of time-dependent systems. This volume will concern itself only with the former part, the problem of equilibrium systems.

There exists a rigorous, concise, and inclusive macroscopic theory for the treatment of systems in equilibrium, namely thermodynamic theory. It is always possible to use statistical mechanical equations in such a way as to circumvent the necessity of computing purely thermodynamically defined quantities, such as entropy, and to proceed to the immediate computation of the results of direct physical measurements. Such a bypass of the classic methods of thermodynamics, seems, however, to be inefficient and inelegant. The more usual procedure, and that which will be pursued here, is to use the method of statistical mechanics to set up equations for the computation of one or other of the thermodynamic functions of a system. This function, then, by the usual manipulations of thermodynamics, may be used to predict the values of directly measurable quantities. Thus, for instance, one sets up an equation by means of which the Helmholtz free energy, $A(N, V, T)$, for a particular chemical species of molecule may be computed for any number N of molecules, at any volume V and temperature T. From this, and the usual relations, $P = -(\partial A/\partial V)_{N,T}$, $S = -(\partial A/\partial T)_{N,V}$, any thermodynamic property of the system is obtainable.

The complete outline of equilibrium statistical mechanics may be broken into three parts. These are as follows:

(1) The derivation of the laws of thermodynamics as a consequence of the laws of mechanics, rather than as a separate axiomatic assumption of science.

(2) The derivation of equations by means of which any of the fundamental thermodynamic functions may be computed, at least in principle, from a knowledge of the microscopic properties of the constituent molecules for any system.

(3) The methods by which, in sufficiently simple cases, the concise, but still complicated equations for the thermodynamic functions may be numerically evaluated.

1.2. A master equation

Rather than proceed in the logical order described at the end of the last section, we choose here to give the answer to the second step, namely to state one of the alternative equations by which the properties of any system may be computed, and then proceed immediately to the third problem of showing how the equation may be used. The reader whose scientific sensitivities are too badly bruised by being presented with an equation without proof may proceed to Chapter 8 where the more usual procedure of formal development is started. Others may find Chapter 8 less abstract if they understand the use and nature of the equations to be derived.

The equation which we give, eq. (4), is itself no more or less fundamental than a number of others of similar nature which can readily be derived from it, and from which it can be derived. It does, however, represent a convenient starting point as a "master equation" from which all results of equilibrium statistical mechanics follow.

The equation refers to what is called the "Grand Canonical Ensemble," and gives the probability that a member of the ensemble will have a certain set, $\mathbf{N}$, of molecules within it, and that these will be in a given quantum state, $\mathbf{K}$.

Consider a reservoir of infinite extent at equilibrium at temperature T, and for which the chemical potentials $\mu_a, \mu_b, \ldots$, etc., of all species $a, b, \ldots$ of molecules composing the system are specified. These intensive thermodynamic variables are sufficient to specify the intensive properties of the material. Use the shorthand notation,

$$\boldsymbol{\mu} = \mu_a, \mu_b, \ldots, \tag{1.2.1}$$

for the set, $\boldsymbol{\mu}$, of chemical potentials. In such an infinite system imagine a finite, but macroscopic volume, V, bounded by mathematically defined planes. Since, later, we wish to discuss the properties of the molecules within V, but to neglect their molecular interactions with molecules outside of V, it is necessary that V be large enough (and of such a shape) that the fraction of molecules at the surface is negligible.

Consider, now, a very large number of such sub-volumes, V, each in equilibrium with the infinite reservoir of fixed T and $\boldsymbol{\mu}$, and each of the same size and shape. Indeed, we imagine the total number of these sub-volumes to approach infinity in value. The grand canonical ensemble consists of the totality of all these sub-volumes. The individual systems in each V are said to be "members" of the ensemble. They are often referred to as "open" systems.

We wish to discuss the probability that a member of this ensemble, selected at random, will have certain properties. In general, since the systems are open, and molecules of the chemical species, a, b, etc., constantly cross the walls of the systems, the numbers, N_a, N_b, . . ., etc., of the different species within different members of the ensemble will fluctuate. Use,

$$\mathbf{N} = N_a, N_b, \ldots, \tag{1.2.2}$$

to indicate the number set within V of a member of the ensemble at some specified time, and the notation,

$$\mathbf{N} \cdot \boldsymbol{\mu} = N_a\mu_a + N_b\mu_b + \ldots. \tag{1.2.3}$$

The probability W ($\mathbf{N}$, $\mathbf{K}$) that a member of the ensemble will, at a given instant, contain exactly the number set $\mathbf{N}$ of molecules, and that this system of N_a, N_b, . . . molecules in the volume V will be in the quantum state, $\mathbf{K}$, of energy E_k, is, with P the pressure in the ensemble,

$$W(\mathbf{N}, \mathbf{K}) = \exp\,[1/kT]\,[-PV + \mathbf{N} \cdot \boldsymbol{\mu} - E_k]. \tag{1.2.4}$$

The sum of $W(\mathbf{N}, \mathbf{K})$ over all $\mathbf{K}$ and all $\mathbf{N}$ is unity, since W is a probability, and this condition fixes P for given V, T, and $\boldsymbol{\mu}$. From this one equation all the results of equilibrium statistical mechanics follow.

The quantity k in kT of eq. (4) is the Boltzmann constant, of dimension energy deg^{-1}. Its numerical value is

$$\begin{aligned} k &= 1.380\,54 \times 10^{-16}\,\text{erg/deg} \\ &= 8.616 \times 10^{-5}\,\text{eV/deg} \end{aligned} \tag{1.2.5}$$

when T is in the Kelvin scale.

If N_0 is Avogadro's number, $N_0 = 6{\cdot}022\,52 \times 10^{23}$

$$R = N_0k, \qquad k = R/N_0, \tag{1.2.6}$$

where R is the usual gas constant per mole, 8.314 3 joule/mol.

Before proceeding with the use of eq. (4) to obtain some results, some comments on the meaning of the "quantum state of the system, $\mathbf{K}$" seem to be required. In the first place the specification of $\mathbf{K}$ requires the specification of at least $3N$ different numbers, namely as many single quantum numbers as the "number of degrees of freedom" of the set N of molecules, which number is equal to the total number of coordinates required to specify exactly the position and internal state of each molecule. It is exactly the purpose of introducing and using eq. (4) to show early, by example, the significance of this concept, and an example in which $\mathbf{K}$ can be readily specified is given in section 1.5.

However, a question of terminology should be disposed of here. In discussing the quantum mechanics of a simple system, such as an electron in a hydrogen atom, or the nuclear motion of a diatomic molecule, one seeks solutions of the time independent Schrödinger equation,

$$\mathscr{H}\psi_n = E_n\psi_n \tag{1.2.7}$$

in which $\mathscr{H}$ is the quantum mechanical Hamiltonian operator, ψ_n is a function of the coordinates of the system, and E_n is an energy value. If the Hamiltonian has some symmetry it frequently happens that several, say g, different mutually orthogonal solutions $\psi_{n,1}, \psi_{n,2}, \cdots, \psi_{n,g}$, are found for a single energy E_n. The quantum "state" of the system specifies the particular functions $\psi_{n,\nu}$ and g is said to be the "degeneracy" of the "level." Here, and in all that follows, we use the term "quantum state" to refer to the completely specified state, independently of whether other states of the same energy may exist, and do not refer to the level which may be degenerate. We may, parenthetically, remark here that we shall later find that the degeneracy of a specified energy E of a macroscopic system has a fantastically large value, of order $10^{10^{20}}$. In other words there are some $10^{10^{20}}$ different $\boldsymbol{K}$'s with the same energy E.

1.3. The average of a mechanical property

The knowledge of the values of the probabilities, $W(\boldsymbol{N}, \boldsymbol{K})$, of eq. (1.2.4) gives answers, as detailed as it is possible to give, to any question about the physical behavior of the systems in the open ensemble.

The classical Hamiltonian $H_{\boldsymbol{N}}(\boldsymbol{p}^{(N)}, \boldsymbol{q}^{(N)})$, for a system of the set $\boldsymbol{N}$ of molecules is a function of all the coordinates,

$$\boldsymbol{q}^{(N)} = q_1, q_2, \cdots, q_\Gamma, \tag{1.3.1}$$

and momenta,

$$\boldsymbol{p}^{(N)} = p_1, p_2, \cdots, p_\Gamma, \tag{1.3.2}$$

required for the complete specification of the mechanical state of the system. The number, Γ, of these single coordinates is known as the "Number of degrees of freedom" of the system, since non-holonomic systems, in which the number of coordinates and degrees of freedom differ, do not occur on a microscopic level. The quantum mechanical Hamiltonian operator, $\mathscr{H}$, is obtained from $H(\boldsymbol{p}^{(N)}, \boldsymbol{q}^{(N)})$ by replacing p_j with $(\hbar/i)\,\partial/\partial q_j$, paying respectful attention to cases where noncommutability of the order p_jq_j or q_jp_j cause difficulty. The time independent Schrödinger equation,

$$\mathscr{H}\psi_{\boldsymbol{K}}(\boldsymbol{q})^{(N)} = E_{\boldsymbol{K}}\psi_{\boldsymbol{K}}(\boldsymbol{q}^{(N)}), \tag{1.3.3}$$

is then used to find the normalized and mutually orthogonal stationary state solutions $\psi_K(\boldsymbol{q}^{(N)})$, subject to suitable boundary and symmetry conditions, each state $\boldsymbol{K}$ having a fixed energy $E_{\boldsymbol{K}}$.

Equation (1.2.4) giving the probability $W(\boldsymbol{N}, \boldsymbol{K})$ that a system of the open ensemble of fixed T and $\boldsymbol{\mu}$ will have the set $\boldsymbol{N}$ of molecules in V in the state $\boldsymbol{K}$ differs only if their energies, E_K, differ. However, the state function $\psi_K(\boldsymbol{q}^{(N)})$ contains the most detailed information that can be given about the mechanical properties of the system in the state $\boldsymbol{K}$. The average value, $<F_K>$, of any analytic function, $F(\boldsymbol{p}^{(N)}, \boldsymbol{q}^{(N)})$, of the momenta and coordinates of the system in state $\boldsymbol{K}$ is obtained by using the operator $\mathscr{F}$, obtained by replacing p_j in F with $(\hbar/i)(\partial/\partial q_j)$, and integrating,

$$<F_K> = \int . . . \int \psi^*{}_K \mathscr{F} \psi_K \, d\boldsymbol{q}^{(N)}. \qquad (1.3.4)$$

The average over the open ensemble is then

$$<F> = \sum_{\boldsymbol{N}}\sum_{\boldsymbol{K}}{}^{(N,V)} \; W(\boldsymbol{N},\boldsymbol{K}) <F_K> \qquad (1.3.5)$$

This rather pompous statement that a physical quantity describable as a function of the mechanical coordinates and momenta of the molecules which means any physically measurable quantity, has an average value $<F>$ which can be computed from the $W(\boldsymbol{N}, \boldsymbol{K})$ of (1.2.4) is inserted here merely to emphasize that eq. (1.2.4) does actually solve, in principle, the problem of determining the properties of the bulk system. The actual computation, following the prescription (4) plus (5) for a complicated system, and a complicated function, $F(\boldsymbol{p}^{(N)}, \boldsymbol{q}^{(N)})$, might require extraordinary mathematical virtuosity. However, several practical questions can actually be answered for the simpler types of systems. We will here confine our attention to two simple quantities, the average energy and the average number of molecules in V and show that eq. (1.2.4) does lead to value consistent with those of thermodynamics.

The operator corresponding to the energy function of $\boldsymbol{p}^{(N)}\boldsymbol{q}^{(N)}$ is the Hamiltonian operator, which from (3) gives $E_K\psi_K$ when operating on ψ_K, with E_K a constant. Since the integral $\psi_K{}^*\psi_K$ over $dq^{(N)}$ is normalized to unity, prescription (4) gives us E_K as the energy of state $\boldsymbol{K}$, which, for the stationary states is not merely an average, but an exact value. The average energy of the ensemble is, from eq. (5),

$$\bar{E} = \sum_{\boldsymbol{N}}\sum_{\boldsymbol{K}}{}^{(N,V)} E_K \; W(\boldsymbol{N},\boldsymbol{K}), \qquad (1.3.6)$$

where the superscript (N,V) on the summation over K is used to remind us that only quantum states of fixed N and V are included.

The average number, $\bar{N}_a$, of molecules of species a in the volume V of the open ensemble is given by

$$\bar{N}_a = \sum_N N_a \sum_K^{(N,V)} W(N,K). \tag{1.3.7}$$

One may write, symbolically, for the average set, $\bar{N} = \bar{N}_a, \bar{N}_b, \ldots,$

$$\bar{N} = \sum_N N \sum_K^{(N,V)} W(N,K) \tag{1.3.8}$$

and for the total Gibbs free energy, G, which is $\bar{N} \cdot \boldsymbol{\mu}$,

$$\bar{G} = \bar{N} \cdot \boldsymbol{\mu} = \sum_N N \cdot \boldsymbol{\mu} \sum_K^{(N,V)} W(N,K). \tag{1.3.9}$$

Using eq. (1.2.4) for $W(N,K)$ one has that

$$(\partial W(N,K)/\partial \mu_a)_{T,\mu_b}\ldots = (kT)^{-1}[N_a - V(\partial P/\partial \mu_a)_{T,\mu_b}\ldots]W(N,K), \tag{1.3.10}$$

and

$$T(\partial W(N,K)/\partial T)_\mu = (kT)^{-1}[PV - N \cdot \boldsymbol{\mu} + E_K - VT(\partial P/\partial T)_\mu]W(N,K). \tag{1.3.11}$$

Now, the sum over N and K of $W(N,K)$ must be identically unity, since the volume V must contain some set N between zero and infinity, and these must be in some state K. The derivatives of the sum with respect to μ_a or T are therefore zero. Using these relations, summing (10) and (11) over N and K one finds, with (6), (7) and (9), that

$$(\partial P/\partial \mu_a)_{T,\mu_b}\ldots = \bar{N}_a/V, \tag{1.3.12}$$

$$(\partial P/\partial T)_\mu = (VT)^{-1}[\bar{E} + PV - \bar{G}] = +\bar{S}/V, \tag{1.3.13}$$

where $\bar{S}$ is the average entropy, $T\bar{S} = \bar{E} + PV - \bar{G}$.

Now on the other hand there are purely thermodynamic equations for these derivatives. One may write

$$-PV = A - G = A - N \cdot \boldsymbol{\mu}, \tag{1.3.14}$$

and since

$$dA = -PdV - SdT + \boldsymbol{\mu} \cdot dN, \tag{1.3.15}$$

$$d(N \cdot \boldsymbol{\mu}) = \boldsymbol{\mu} \cdot dN + N \cdot d\mu, \tag{1.3.16}$$

one has

$$d(-PV) = -PdV - SdT - N \cdot d\mu \tag{1.3.17}$$

or, since P does not depend on V at constant $T,\boldsymbol{\mu}$,

$$(\partial P/\partial \mu_a)_{T,\mu_b}\ldots = N_a/V, \tag{1.3.18}$$

$$(\partial P/\partial T)_{\boldsymbol{\mu}} = S/V \tag{1.3.19}$$

where the subscript V has been omitted from the derivatives since P depends only on the intensive variable set $\boldsymbol{\mu}, T$.

The thermodynamic equations (18) and (19) agree with (12) and (13) which were derived purely from the interpretation of $W(\boldsymbol{N},\boldsymbol{K})$ as a probability, and the rules for finding averages. The expression (1.2.4) for $W(\boldsymbol{N},\boldsymbol{K})$ is thus consistent with the laws of thermodynamics.

However, one difference in the equations (18), (19), and (12), (13) is apparent. In classical thermodynamics one assumes that if the state of the system is determined, for instance by the variables V, T, and $\boldsymbol{\mu}$, then the other thermodynamic quantities such as $P, \boldsymbol{N} = N_a N_b, \ldots,$ E, S, and G, etc., are uniquely determined. In our statistical derivations, however, the ensemble was defined by the values of V, T, and $\boldsymbol{\mu}$ but the energy $\bar{E}$, and the numbers of molecules, $\bar{N}_a$ of species a, and hence also $\bar{G}$ and $\bar{S}$ were computed only as *average* values. Since we know empirically, and can later use our equations to predict, that fluctuations from the average, large enough to be experimentally observable, simply do not occur with any appreciable probability in large systems, we will identify the "average values" with the exact values predicted from thermodynamic theory.

In the next section, we shall have occasion to derive a relationship by identifying the *most probable* number set, $\boldsymbol{N}^*$, with the thermodynamic value of $\boldsymbol{N}$. In general $\bar{\boldsymbol{N}}$ and $\boldsymbol{N}^*$ are not identical, but if the probability function is sufficiently peaked, that is if detectable percentual fluctuations are not observed, the two sets will be numerically indistinguishable. The identification of $\boldsymbol{N}^*$ with $\bar{\boldsymbol{N}}$, and both with the thermodynamic $\boldsymbol{N}$ depends critically on the fact that we implicitly limit our considerations to large systems. Only for large systems can we consider the thermodynamic relations to be strictly correct.

1.4. The grand canonical and canonical partition functions

The sum of the probabilities $W(\boldsymbol{N}, \boldsymbol{K})$ over all possible quantum states $\boldsymbol{K}$ of a system with number set $\boldsymbol{N}$, and over all possible number sets $\boldsymbol{N}$ must be unity.

$$\sum_{\boldsymbol{N}}\sum_{\boldsymbol{K}}^{(\boldsymbol{N},V)} W(\boldsymbol{N}, \boldsymbol{K}) = 1 \tag{1.4.1}$$

The term $e^{-PV/kT}$ is a constant factor of each W. We can divide both sides of eq. (1) by this factor. Define

$$Q_{G\cdot C}(V,T,\mu) = \sum_{N\geqslant 0} e^{N\cdot\mu/kT} \sum_{K}^{(N,V)} e^{-E_K/kT}, \tag{1.4.2}$$

and one has, from (1), with (1.2.4), that

$$PV = kT \ln Q_{G\cdot C}. \tag{1.4.3}$$

The dimensionless function $Q_{G\cdot C}$, of T and μ is known as the partition function of the grand canonical ensemble, or more simply the grand canonical partition function.

The quantum states $\boldsymbol{K}$ and their energies E_K, in the sum over $\boldsymbol{K}$ of eq. (2) depend on the set $\boldsymbol{N}$ of molecules, and on the volume V of the system. The function,

$$Q_C(V,T,\boldsymbol{N}) = \sum_{K}^{(N,V)} e^{-E_K/kT}, \tag{1.4.4}$$

is known as the partition function of the macrocanonical ensemble, the "Petite Canonical Ensemble" of Gibbs, or sometimes just as the canonical partition function. These two partition functions are then related from eq. (2) by

$$Q_{G\cdot C}(V,T,\mu) = \sum_{N>0} e^{N\cdot\mu/kT} Q_C(V,T,\boldsymbol{N}). \tag{1.4.5}$$

From (1.2.4) and the fact that the sum of $W(\boldsymbol{N},\boldsymbol{K})$ over $\boldsymbol{K}$ with fixed $\boldsymbol{N}$ gives the probability, $W(\boldsymbol{N})$, of finding exactly the set $\boldsymbol{N}$ of molecules in V, it is seen that the term $\boldsymbol{N}$ in this sum is proportional to $W(\boldsymbol{N})$.

The partition functions for macroscopic systems are extremely large numbers, and this fact is very essential for the validity of many of the manipulations made in statistical mechanics. The fact that $Q_{G\cdot C}$ is large becomes pragmatically evident if we remember that PV/kT for a gas is equal to $\bar{N}$, the total average number of molecules in the system. This means that $Q_{G\cdot C}$ is of the order of $e^{\bar{N}}$, a number of considerably larger magnitude than that encountered in other fields of scientific theory. It is larger by far, when $\bar{N}$ is Avogadro's number, than the number of atoms in the finite universe.

The terms in the sum of eq. (5) are all positive. The sum is therefore larger than any single term. On the other hand it must be less than the value of the largest term multiplied by the total number of terms. The logarithm of the sum can therefore be bracketed in value between the value of the logarithm of the largest term and the value of this plus the logarithm of the total number of terms. In a hypothetical case, if we have

$$\sigma = \sum_{n=0}^{n=M} T_n, \qquad T_n \geqslant 0, \tag{1.4.6}$$

then, if m is the value of n for which T_n is a maximum,

$$T_m \geqslant T_n, \qquad \text{all } n \neq m, \tag{1.4.7}$$

$$\ln T_m \leqslant \ln \sigma \leqslant \ln T_m[1 + \ln M/\ln T_m]. \tag{1.4.8}$$

Now if, as is usually the case in statistical mechanical sums of this nature, the value of $\ln T_m$ is given by

$$\ln T_m = \nu^{-1}M, \tag{1.4.9}$$

with ν of order unity, and M is of order 10^{20} or more, $\ln M = 46$, then,

$$\ln T_M \leqslant \ln \sigma \leqslant \ln T_M(1 + \nu M^{-1}\ln M) = \ln T_M(1 + 46\nu 10^{-20}). \tag{1.4.10}$$

We may therefore write $\ln \sigma = \ln T_m$ with a very high degree of precision.

This argument has been given in some detail here, since the replacement of the logarithm of a sum of positive terms by the logarithm of the largest term will be used repeatedly. As implied earlier, the numerical values corresponding to M are generally proportional to the size of the system, usually of the order of the average number of molecules. Although for ordinary macroscopic systems with $\bar{N}$ of order 10^{20} or greater, the replacement of $\ln T_m$ is numerically valid, this would not be the case were we dealing with systems of only a few hundred molecules. In these cases of extremely minute systems the failure of the replacements to be valid reflects a necessity to modify the usual thermodynamic identities by considering the possibilities of large fluctuations from mean values.

The application of the argument above to eq. (5) is not as clean as in the example given since the sum over N ostensibly runs to infinity. However, real molecules repel each other, and only a finite number, M, can occupy any finite volume V, where $M = V/r_0^3$ and r_0 is some kind of minimum distance of approach.

If the molecules did not have this property, that is if the forces were purely attractive, and sums of the pair forces, so that any number could occupy a finite volume, the potential energy would be proportional to the number of pairs, namely $(1/2)N(N - 1)$, and not to the number of molecules. The energy would not be proportional to N, but to N^2, and hence not an extensive property. The normal thermodynamic equations would not apply.

Since the terms of eq. (5) are proportional to the probabilities of having the set N in V all terms will be zero for $N = \sum_a N_a > M$, and we can regard the sum of eq. (5) as finite. With j chemical components the order of the number of terms would be M^j and eq. (10) would be applicable with an extra j as factor in the correction term.

We may therefore write from (2) and (3)

$$PV = kT \ln Q_{G\cdot C} = \mathbf{N}^* \cdot \boldsymbol{\mu} + kT \ln Q_C(V,T,\mathbf{N}^*), \qquad (1.4.11)$$

where $\mathbf{N}^*$ is used to indicate that number set $\mathbf{N}$ of molecules which makes $\mathbf{N} \cdot \boldsymbol{\mu} + kT \ln Q_C$ a maximum. This term, apart from the additive constant $-PV$, is kT times the logarithm of the probability of finding the set $\mathbf{N}^*$ of molecules in V in the open system of the grand canonical ensemble. The value of $\mathbf{N}^*$ is that of *the most probable number set of molecules* which will occupy the volume V, when T and the chemical potential set $\boldsymbol{\mu}$ are specified.

Since $\mathbf{N} \cdot \boldsymbol{\mu} = G$, the Gibbs free energy, and $G - PV = A$, the Helmholtz free energy, we have

$$A(V,T,N) = -kT \ln Q_C(V,T,N), \qquad (1.4.12)$$

where we identify the most probable number set $\mathbf{N}^*$ with the actual number set N in a closed system for which the numbers of molecules are held fixed.

The value of $\mathbf{N}^*$ to be used in (11), since it is that for which the term is a maximum, is that for which the derivative with respect to N_a for every a is zero,

$$[(\partial/\partial N_a)(\mathbf{N} \cdot \boldsymbol{\mu} + kT \ln Q_C(V,T,\mathbf{N})]_{\mathbf{N}=\mathbf{N}^*} = 0. \qquad (1.4.13)$$

Using (12) for $kT \ln Q_C$ this gives

$$\partial A(V,T,N)/\partial N_a = \mu_a, \qquad (1.4.14)$$

again consistent with the thermodynamic relations.

In thermodynamic theory there exist natural extensive energy quantities to various sets of the variables defining a system. These energy quantities are those that have minimum value at equilibrium for any virtual displacement of the system at fixed values of their natural variables, and the partial derivatives of the energy quantity with respect to its natural variables are, with plus or minus signs, the conjugate variables. Thus the Helmholtz free energy, A, has natural variables, V, T, N, and its partial derivatives with respect to these variables are $-P$, $-S$, and μ, respectively. In the last section we saw that the extensive energy function $-PV$ had natural variables V, T, and μ, since its derivatives with respect to these were $-P$, $-S$, and $-N$, respectively. The quantity $-PV$ is a minimum at equilibrium if V, T, μ are fixed.

Both quantities, $-PV$ and A, are given by eqs. (3) and (12) as equal to $-KT$ times the logarithm of a partition function, $Q_{G \cdot C}$ and Q_C respectively. The ensemble for which $Q_{G \cdot C}$ is computed is that in which the variables V, T, μ are held fixed, the sum runs over all energies and all N. The ensemble for which Q_C is computed is that characterized by fixed values of V, T, and N, the sum running over all energies. In section 8.9 we shall discuss other partition functions. In general whenever the partition function is that of an ensemble of fixed intensive variable the partition function is a sum, or integral, over the conjugate extensive variable. The logarithm of the partition function, times $-kT$, is equal to the natural energy function of the variables fixed in the ensemble.

1.5. Perfect gases

Equations (1.4.2) and (1.4.3) may be used readily to obtain the thermodynamic equations for a perfect gas. Their use illustrates the significance of the identification symbol K for the specification of the quantum state of the system in this particularly simple case. To simplify the problem we limit the discussion to the example of a pure one component system for which only one chemical species has μ differing from minus infinity, so that non-zero terms in the sum (1.4.2) occur only with one $N_a = N$ differing from zero. We have then that

$$PV = kT \ln \sum_{N=0} e^{N\mu/kT} \sum_{K}^{(N,V)} e^{-E_K/kT}. \tag{1.5.1}$$

For a perfect gas, which we define mechanically be the statement that there are no forces between different molecules, the molecules are independent, and each molecule may be assigned a quantum number $\boldsymbol{m}$ describing its state in the volume V. This quantum number $\boldsymbol{m}$, is itself a number which has at least three components. There are three quantum numbers k_x, k_y, k_z, of the translational motion of the center of mass coordinates x, y, z, in the state

$$\psi_k(xyz) = \sqrt{(8/L_xL_yL_z)} \sin(\pi k_x x/L_x) \sin(\pi k_y y/L_y \sin(\pi k_z z/L_z) \tag{1.5.2}$$

which, for integer values of k_x, k_y, k_z, are the stationary states for a rectangular box of edge lengths L_x, L_y, L_z and volume $V = L_xL_yL_z$. In addition there will generally be an internal quantum number, $\boldsymbol{n}$, having components which can be identified with those describing the electronic, vibrational and rotational states of the molecule. The energy of the center of mass motion is the kinetic energy,

$$\varepsilon_k = (h^2/2m)[(k_x/2L_x)^2 + (k_y/2L_y)^2 + (k_z/2L_z)^2], \tag{1.5.3}$$

In this, as also later we use the symbol ε to indicate energies of single molecules, with lower case letters k, m, etc., to indicate the quantum states. Capital letters $\boldsymbol{K}$, etc., are used for quantum states of systems of many molecules and E_K for their energies with m the molecular mass. If we use the symbol ε_n for the internal energy in the quantum state $\boldsymbol{n}$, the total energy of the molecule in state $\boldsymbol{m}$ is

$$\varepsilon_m = \varepsilon_k + \varepsilon_n \tag{1.5.4}$$

Now the complete quantum state $\boldsymbol{K}$ of the system is given by specifying the numbers, n_m, of identical molecules (the case of isotopes is discussed in section 2.6) having the quantum state $\boldsymbol{m}$, for all $\boldsymbol{m}$,

$$\boldsymbol{K} = n_0, \ldots, n_m, \ldots, \tag{1.5.5}$$

and the energy E_K is the sum of the energies of the independent molecules,

$$E_K = \sum_m n_m \varepsilon_m. \tag{1.5.6}$$

For given value, N, of the numbers of molecules in V one must have

$$\sum_m n_m = N, \tag{1.5.7}$$

since every molecule is in some state $\boldsymbol{m}$. One may therefore write,

$$N\mu = \sum_m n_m \mu, \tag{1.5.8}$$

and,

$$N\mu - E_K = \sum_m n_m(\mu - \varepsilon_m). \tag{1.5.9}$$

The quantity $(N\mu - E_K)/kT$ occurs in the exponent of (1). The exponential of a sum is a product. Summing over all values of n_m consistent with (7), and overall values of N from zero to infinity means that we may sum over all values of n_m for each $\boldsymbol{m}$ independently, thus finding the sum under the logarithm of (1) to be a product of sums. The logarithm of a product is the sum of the logarithms, so that (1) becomes,

$$PV = kT \sum_m \ln \sum_n e^{n(\mu - \varepsilon_m)/kT}. \tag{1.5.10}$$

Now two cases occur, namely the molecules may form a Bose–Einstein or a Fermi–Dirac system, depending on whether the single molecule consists of an even or odd number of the traditional fundamental particles respectively. If the sum of the numbers of neutrons, protons, and electrons in the molecule is even the total state function of the system is symmetric in permutations of the identical molecules, and functions in which two ore more molecules are in identical quantum states $\boldsymbol{m}$ exist and are allowed. The system is said to be a Bose–Einstein system, or frequently said to obey Bose–Einstein statistics. In this case the sum in (10) over n goes from zero to infinity. If the sum of fundamental particles is odd, only antisymmetric state functions of the system occur, and these are identically zero if there is more than one molecule in any single molecular quantum state $\boldsymbol{m}$. The system is said to be a Fermi–Dirac system, and the sum over n in (10) includes only the terms $n = 0$ and $n = 1$.

Since for the Bose–Einstein case we may use,

$$\sum_{n=0} x^n = (1 - x)^{-1}, \quad x < 1,$$

we have from (10), for $\mu - \varepsilon_k < 0$,

$$PV = kT \sum_{m} \mp \ln[1 \mp e^{(\mu - \varepsilon_m)/kT}] \begin{Bmatrix} -\text{B.E.} \\ +\text{F.D.} \end{Bmatrix}. \tag{1.5.11}$$

The chemical potential μ, if measured from the zero of energy of the lowest internal quantum state of the molecule such that $\varepsilon_m \geqslant 0$, $\varepsilon_0 = 0$, is negative at high temperatures, and indeed for all other molecular species than helium the numerical value of μ/kT is negative and rather large in magnitude whenever the material exists as a gas. The exponential $e^{\mu/kT}$ is hence small compared with unity, and $\mp \ln(1 \mp x) = x$ for $x << 1$. The classical limit of what is referred to as "Boltzmann Statistics" is then obtained

$$PV = kT \sum_{m} e^{(\mu - \varepsilon_m)/kT}, \tag{1.5.12}$$

The sum over the quantum states $\boldsymbol{m}$ goes independently over the set $\boldsymbol{k}$ of positive translational quantum numbers, and over the set $\boldsymbol{n}$ of internal quantum numbers. The energy ε_m is a sum, from (4), so that the exponential is a product, and the sum over m a product of two sums,

$$Q_{\text{tr}}(V,T) = \sum_{k} e^{-\varepsilon_k/kT}, \tag{1.5.13}$$

$$Q_i(T) \quad = \sum_{n} e^{-\varepsilon_n/kT}, \tag{1.5.14}$$

and (12) becomes

$$\mu = kT \ln (PV/kT\, Q_{\text{tr}} Q_i). \tag{1.5.15}$$

The sum Q_{tr} is evaluated as a classical integral over momentum space. The stationary state (2), for a positive set k_x, k_y, k_z, is a superposition of progressive waves in which the molecule moves with magnitudes $p_x = hk_x/2L_x$, $p_y = hk_y/2L_y$, . . ., of the momentum vector, but with both positive and negative values of each component. The energy (3) is then the classical kinetic energy,

$$\varepsilon_k = \varepsilon(p) = (1/2)p^2/m, \qquad p^2 = p_x{}^2 + p_y{}^2 + p_z{}^2. \tag{1.5.16}$$

For macroscopic values of the L's, the consecutive states lie so closely together in energy compared with kT that summation over k can be replaced by integration over the positive octant of $dk_x\, dk_y\, dk_z$. Divide by 8, and integrate over k-space from $-\infty$ to $+\infty$,

$$\begin{aligned}\iint\int_0^\infty dk_x\, dk_y\, dk_z &= (1/8)\iint\int_{-\infty}^{+\infty} dk_x\, dk_y\, dk_z \\ &= (L_x L_y L_z/h^3)\iint\int_{-0}^{+\infty} dp_x\, dp_y\, dp_z = 4\pi V/h^3 \int_0^\infty p^2 dp.\end{aligned} \tag{1.5.17}$$

One has

$$\begin{aligned}Q_{\text{tr}} &= (4\pi V/h^3)\int_0^\infty p^2 e^{-p^2/2mkT} dp \\ &= V(2\pi mkT/h^2)^{3/2}.\end{aligned} \tag{1.5.18}$$

Using this in (15) we find

$$\mu = kT \ln (P/kT)(h^2/2\pi mkT)^{3/2} Q_i^{-1}. \tag{1.5.19}$$

for the chemical potential, μ, as a function of P and T. The thermodynamic relation for the volume, v, per molecule, since Q_i is independent of V, is

$$V/\bar{N} = v = (\partial\mu/\partial P)_T = kT/P, \quad P = kT/v, \tag{1.5.20}$$

gives the perfect gas law.

The quantity Q_i of eq. (14) is known as the internal partition function of the molecule, and Q_{tr} of (13) is often called the translational partition function of the molecule. Whereas the form of Q_{tr} is identical for all independent (perfect gas) molecules, depending on the molecular properties only through the total mass, m, by (18), the internal partition function Q_i depends on the details of the molecular structure. Its evaluation for various cases is discussed in Chapter 3. We wish, however, to emphasize here only that Q_i, unlike the system partition functions $Q_{G \cdot C}$ and Q_C, has comparatively small numerical values. Indeed Q_i is effectively unity for the noble gases at all normal temperatures, since the only internal excited states are electronic, and these have such high energies compared with kT that only the lowest single state of energy $\varepsilon_0 = 0$, $e^{\;\varepsilon/kT} = 1$, has non-negligible value in the sum (14).

One other remark is in order at this place. The zero of energy chosen for a particular molecular species is arbitrary. We have implicitly assumed, in the foregoing, that the lowest internal quantum state of the molecule is assigned zero energy. This means that the thermodynamic zero of energy is chosen as that of all molecules at rest in their lowest quantum state, but in the perfect gas state of zero (or negligible) mutual interaction potential. In other words the energy zero is that of the perfect gas extrapolated to $T = 0°K$. With this choice, the value of μ computed from (19) will be that consistent with this zero. If the lowest internal molecular state is assigned an arbitrary energy ε, then from (14) Q_i is multiplied by the factor $e^{-\varepsilon/kT}$, and from (19) μ is raised by the additive term ε, which is consistent.

In the more general case of the master equation (1.2.4) the energy $E_{\boldsymbol{K}}$ of the system in quantum state $\boldsymbol{K}$ would be raised by $N_a\varepsilon_a$ if the arbitrary zero of energy per molecule of species a were raised by ε_a. However $\boldsymbol{N} \cdot \boldsymbol{\mu}$ would be raised by the same amount, and the probability expression (1.2.4) would remain unchanged in value. The equations are independent of the choice of the energy zero, as they should be, but, of course, consistency is required.

The quantity $(h^2/2\pi mkT)^{1/2}$ in (18) has dimension length. It occurs so frequently in equations that it is convenient to assign to it a special symbol,

$$\lambda = (h^2/2\pi mkT)^{1/2}. \tag{1.5.21}$$

Numerically one finds, with M the molecular weight, that

$$\left.\begin{aligned} \lambda &= A \times 10^{-7}\, M^{-1/2}\, T^{-1/2}, \\ A &= 1.746\ 11 \text{ cm deg}^{1/2}, \\ A^2 &= 3.048\ 91 \text{ cm}^2 \text{ deg} \\ A^3 &= 5.323\ 73 \text{ cm}^3 \text{ deg}^{3/2}. \end{aligned}\right\} \tag{1.5.22}$$

A convenient method of remembering is that

$$\lambda = \sqrt{(300/MT)} \text{ angstrom units}, \tag{1.5.23}$$

within about one per cent.

The numerical value of $e^{\mu/kT}$ from (19) is

$$\left.\begin{aligned} e^{\mu/kT} &= \frac{P\lambda^3}{kTQ_i} = BPM^{-3/2}\,T^{-5/2}Q_i^{-1} = B'PM^{-3/2}(300/T)^{5/2}Q_i^{-1} \\ B &= 39.078 \text{ atm}^{-1}\text{ deg}^{5/2},\ B' = 2.5069\ 10^{-5}\text{ atm}^{-1}\text{ deg}^{5/2} \end{aligned}\right\} \quad (1.5.24)$$

Since Q_i is unity or greater we see that at room temperature, $T = 300°K$, the value of $e^{\mu/kT}$ is indeed small, namely of order $10^{-5}M^{-3/2}$, and the assumption that leads to the Boltzmann equation (12) is numerically justified.

The length λ is frequently referred to as the de Broglie wavelength of the molecule. Actually it is one-half of the mean magnitude of wavelength for molecules of mass m at temperature T.

The probability density $W(\boldsymbol{p})$ in momentum space, with $p = |\boldsymbol{p}|$, is

$$W(p) = (2\pi mkT)^{-3/2} \exp\left(-p^2/2mkT\right), \qquad (1.5.25)$$

correctly normalized so that,

$$\iiint_{-\infty}^{+\infty} W(\boldsymbol{p})d\boldsymbol{p} = \int_0^\infty 4\pi p^2 W(\boldsymbol{p})dp = \frac{4}{\sqrt{\pi}}\int_0^\infty y^2 e^{-y^2}\,dy = 1,$$

with $y = p/(2mkT)^{1/2}$.

The de Broglie wavelength is $\Lambda = h/p = \sqrt{\pi}\lambda y^{-1}$, and its average value is then

$$<\Lambda> = <h/p> = h(2\pi mkT)^{-1/2}\,2\int_0^\infty 2ye^{-y^2}\,dy = 2\lambda. \quad (1.5.26)$$

1.6. The occupancy of a molecular quantum state

Our method was chosen as the most rapid and economical to obtain the perfect gas equation. It has, however, served to disguise a very important consequence of eq. (1.2.4). Use (1.5.9) for $N\mu - E_K$ in (1.2.4) to write

$$\begin{aligned} W(N,K) &= \exp\frac{1}{kT}\left[-PV + \sum_m n_m(\mu - \varepsilon_m)\right] \\ &= e^{-PV/kT}\prod_m e^{n_m(\mu-\varepsilon_m)/kT}. \end{aligned} \qquad (1.6.1)$$

The occupancy numbers, n_m, giving the numbers of molecules in the molecular quantum state m, for all m determine both $N = \Sigma n_m$ and the quantum state K. The sum over all allowed values of n_m for each m of (1) is unity. The average number, $\bar{n}_\lambda$, of molecules in state λ is

$$\begin{aligned}\bar{n} &= \sum_{n_0} \sum_{n_1} \cdot\,\cdot\,\cdot \sum_{n_m} \cdot\,\cdot\,\cdot\, n_\lambda\; W(N,K) \\ &= \sum_{n} n e^{n(\mu-\varepsilon_\lambda)/kT} \Big/ \sum_{n} e^{n(\mu-\varepsilon_\lambda)/kT}. \qquad (1.6.2)\end{aligned}$$

We have now the two cases. Bose–Einstein and Fermi–Dirac. In the former case $0 \leqslant n \leqslant \infty$, and $\sum_{n\geqslant 0} x^n = (1-x)^{-1}$ whereas $\Sigma\, nx^n = x(d/dx)\, \Sigma\, x^n = d/dx(1-x)^{-1} = x/(1-x)^2$. One has, then

$$\bar{n}_\lambda = \frac{1}{e^{(\varepsilon_\lambda-\mu)kT} - 1}, \quad \text{(B.E.)}. \qquad (1.6.3$$

In the F.D. case for which $n = 0$ or 1 only (2) gives

$$\bar{n}_\lambda = \frac{1}{e^{(\varepsilon_\lambda-\mu)/kT} + 1} \quad \text{(F.D.)}. \qquad (1.6.4)$$

In both cases, if $-\mu/kT >> 1$ we can neglect the unity in the denominator. One then has the Boltzmann limit,

$$\bar{n}_\lambda = e^{(\mu-\varepsilon_\lambda)/kT} \quad \text{(Boltzmann)}. \qquad (1.6.5)$$

The average number of molecules in the lowest energy state $\lambda = 0,0$ o zero energy, $\varepsilon_{00} = 0$, with *also* zero kinetic energy of translation, is

$$\bar{n}_{00} = e^{\mu/kT} \quad \text{(zero translational and internal state, Boltzmann)} \quad (1.6.6)$$

and

$$\bar{n}_\lambda = \bar{n}_{00} e^{-\varepsilon_\lambda/kT} \quad (\text{Boltzmann, } \varepsilon_0 = 0). \qquad (1.6.7)$$

The number of molecules in a given quantum state is (in the Boltzmann limit) proportional to $\exp(-\varepsilon/kT)$. If we compute the average number $\bar{n}_n$ in a given *internal* quantum state n we must sum over all translational states, which gives a factor Q_{tr}. From (1.5.18) with (1.5.21) for λ one has

$$\bar{n}_0 = V\lambda^{-3} e^{\mu/kT} \quad \left[\begin{array}{c}\text{for the}\\ \text{lowest internal state}\\ \text{(Boltzmann)}\end{array}\right]. \qquad (1.6.8)$$

As $T \to 0$ and $\mu \to 0$ the number $\bar{n}_{00}$ of (6) could be large. However, since the Boltzmann limit is only valid for $-\mu/kT >> 1$, we see that necessarily in the Boltzmann case the average number of molecules *per completely specified state* is always small, if it is not small; the unities in the denominators of (3) or (4) cannot be neglected. The reason that the number per state of (3), (4), (5), or (6) is independent of the total size of the system is simply that the number of translational states in a given momentum or energy range increases linearly with the size of the system, as does the total number of molecules.

1.7. *PV*–energy relationship for monatomic perfect gases

If a perfect gas has no internal energy, that is if there are no internal quantum states of energy different from zero, so that Q_i is the number of states, g, of zero energy (see section 3.1) and independent of temperature, then PV is always proportional to the energy, E. For the usual gas of non-relativistic molecules in a three-dimensional volume the relationship is

$$E = \frac{3}{2} PV. \tag{1.7.1}$$

We shall show that this relationship holds for both perfect Bose–Einstein and for perfect Fermi–Dirac gases, as well as for the common limiting Boltzmann gas.

Firstly we derive a thermodynamic relationship. Introduce

$$\zeta = e^{\mu/kT} \tag{1.7.2}$$

which is a dimensionless activity. Since

$$d \ln \zeta = d\left(\frac{\mu}{kT}\right) = \frac{1}{kT}\,[d\mu - \mu d \ln T],$$

$$d\mu = kT\, d \ln \zeta + \mu d \ln T$$

we may write (1.3.17) for a single component system as

$$d(PV) = PdV + (TS + N\mu) d \ln T + NkT\, d \ln \zeta, \tag{1.7.3}$$

or, since $N\mu = G$, $TS + G = H = E + PV$,

$$[\partial(PV)/\partial \ln T]_{V,\zeta} = E + PV. \tag{1.7.4}$$

Now eq. (1.5.17) gives the number, $C_{\text{tr}}(V,p)dp$, of translational quantum states between magnitude p of momentum and $p + dp$ as

$$C_{\text{tr}}(V,p)dp = V(4\pi/h^3)p^2 dp. \tag{1.7.5}$$

With kinetic energy $\varepsilon = p^2/2m$, $p^2 = 2m\varepsilon$ and $dp = \frac{1}{2}(2m)^{1/2}\, \varepsilon^{-1/2}\, d\varepsilon$ one has for $C_{\text{tr}}(V,\varepsilon)d\varepsilon$ the number of states between energy ε and $\varepsilon + d\varepsilon$

$$C_{\text{tr}}(V,\varepsilon)d\varepsilon = 2\pi V(2m)^{3/2}h^{-3}\, \varepsilon^{1/2} d\varepsilon. \tag{1.7.6}$$

If we introduce

$$x = \varepsilon/kT, \tag{1.7.7}$$

and

$$C_0 = g(2/\sqrt{\pi})(2\pi m/h^2)^{3/2} \tag{1.7.8}$$

where the integer g is the number of internal states of zero energy, which may exceed unity, then the total number of states in the range between x and $x + dx$ is

$$C_{\text{total}}\,(V,x)dx = V(kT)^{3/2}C_0 x^{1/2}dx. \tag{1.7.9}$$

Now in eq. (1.5.11) we may replace the summation over all states m by integration over x from zero to infinity of $C_{\text{total}}\,(V,x)dx$, which with (7) for x and (2) for ζ gives

$$\begin{aligned} PV &= V(kT)^{5/2}C_0 \int_0^\infty x^{1/2}dx(\mp)\ln[1 \mp \zeta e^{-x}] \\ &= V(kT)^{5/2}C_0 f(\zeta) \end{aligned} \tag{1.7.10}$$

where $f(\zeta)$ is the definite integral over x, and depends only on ζ, that is only on μ/kT. We immediately have

$$PV)/\partial \ln T]_{v,\zeta} = \frac{5}{2} PV \tag{1.7.11}$$

and from (4) that this is $E + PV$ one has the relationship $E = \frac{3}{2}PV$ of equation (1).

In the Boltzmann limit,

$$\mp \ln (1 \mp \zeta e^{-x}) = \zeta e^{-x},$$

with

$$\int_0^\infty x^{1/2}\, e^{-x}\, dx = 2/\sqrt{\pi}$$

we have

$$f_{\text{Boltz}}\,(\zeta) = \frac{2}{\sqrt{\pi}}\,\zeta$$

$$P = (kT)^{5/2}\,\frac{2}{\sqrt{\pi}}\,\zeta C_0 = kT\, g\lambda^{-3}\zeta$$

in agreement with (19) where $g = Q_i$.

The proof used depended on two relations, firstly that $C(V,p)dp \sim Vp^2dp$, and secondly that the energy ε was related to p by $\varepsilon \sim p^2$ so that $C(V,\varepsilon)d\varepsilon \sim V\varepsilon^{1/2}d\varepsilon$. Dividing and multiplying by T to introduce the dimensionless $x = \varepsilon/kT$ one had $P \sim T^{5/2}$ at constant ζ. The first relation, $C(V,p) \sim Vp^2dp$ depends only on the number of dimensions of the volume V, and would be $C \sim Vp^{n-1}dp$ for a hypothetical n-dimensional volume. This relation is valid even if the particles are relativistic, but the relation $\varepsilon \sim p^2$ depends on non-relativistic motion, which, of course, is completely valid for normal molecules at ordinary temperatures. However, one completely relativistic case is of interest. This case is that of light photons (section 7.3) which move with light velocity c, and for which $\varepsilon = pc$ so that $C(V,\varepsilon)d\varepsilon \sim V\varepsilon^2d\varepsilon$. Dividing and multiplying by (kT^3) to convert $x = \varepsilon/kT$ now gives us $P \sim T^4$ at constant ζ. We then have

$$[\partial \ln (PV)/\partial \ln T]_\zeta = 4PV = E + PV \text{ and } E = 3PV.$$

Before proceeding to a further discussion of the evaluation of Q_i and the methods of numerical computation for a perfect gas of a particular molecular species, we shall, in the next chapter, discuss some of the more general characteristics of the manipulations required in the evaluation of $Q_{G \cdot C}$. Many of these characteristics have been encountered by example in this section on the perfect gas.

CHAPTER 2

GENERAL CHARACTERISTICS OF THE PARTITION FUNCTION

2.1. Separability of the Hamiltonian

There is a well-known theorem that if the Hamiltonian of a system is separable in subsets of coordinates and momenta the stationary state functions are products, and the energies sums, of terms dependent on the subset coordinates separately. In this case the partition functions are, in the Boltzmann limit, products of sums, and their logarithms become sums.

If the total set $\boldsymbol{q}^{(N)}$, $\boldsymbol{p}^{(N)}$ of the coordinates and momenta of the system can be decomposed into subsets

$$\boldsymbol{q}^{(N)} = \boldsymbol{q}_\alpha, \quad \boldsymbol{q}_\beta, \ldots, \tag{2.1.1}$$

$$\boldsymbol{p}^{(N)} = \boldsymbol{p}_\alpha, \quad \boldsymbol{p}_\beta, \ldots, \tag{2.1.2}$$

such that one may write,

$$H(\boldsymbol{p}^{(N)}, \boldsymbol{q}^{(N)}) = \sum_\alpha H_\alpha(\boldsymbol{p}_\alpha, \boldsymbol{q}_\alpha), \tag{2.1.3}$$

then, of course, the quantum mechanical operator is also to be written as a sum,

$$\mathscr{H}^{(N)} = \sum_\alpha \mathscr{H}_\alpha. \tag{2.1.4}$$

The solutions, for each α, to the equations,

$$\mathscr{H}_\alpha \psi_{m\alpha}(\boldsymbol{q}_\alpha) = \varepsilon_{m\alpha} \psi_{m\alpha}(\boldsymbol{q}_\alpha) \tag{2.1.5}$$

form a complete set, and the solutions to (1.3.3) may always be written as a product,

$$\psi_{K\dagger}(\boldsymbol{q}^{(N)}) = \prod_\alpha \psi_{m\alpha}(\boldsymbol{q}_\alpha), \tag{2.1.6}$$

$$K\dagger = \boldsymbol{m}_\alpha, \boldsymbol{m}_\beta, \ldots, \tag{2.1.7}$$

$$E_{K\dagger} = \sum_\alpha \varepsilon_{m\alpha}. \tag{2.1.8}$$

The exponential is then a product,

$$\exp -\frac{E_{K\dagger}}{kT} = \prod_{\alpha} e^{-\varepsilon_{m\alpha}/kT}, \tag{2.1.9}$$

and if the sum over $K\dagger$ consists of the independent sum over all quantum numbers m_α for each α,

$$\sum_{K\dagger}\left[\exp -\frac{E_{K\dagger}}{kT}\right] = \sum_{m\alpha}\sum_{m\beta}\cdot\cdot\cdot\prod_{\alpha}\exp -\frac{\varepsilon_{m\alpha}}{kT}$$

$$= \prod_{\alpha}\left(\sum_{m\alpha}\exp -\frac{\varepsilon_{m\alpha}}{kT}\right) = \prod_{\alpha} Q_\alpha, \tag{2.1.10}$$

where

$$Q_\alpha = \sum_{m\alpha}\exp -\frac{\varepsilon_{m\alpha}}{kT} \tag{2.1.11}$$

$$\ln\sum_{K\dagger}\left[\exp -\frac{E_{K\dagger}}{kT}\right] = \sum_{\alpha}\ln Q_\alpha. \tag{2.1.12}$$

We have encountered an example of (3) in the perfect gas. By its mechanical definition as a gas of molecules which have no potential of mutual interaction, the Hamiltonian of the system is a sum of the Hamiltonians of the individual molecules. The α's of (1) are then the indices 1, 2, 3 . . ., N identifying the individual molecules, and the q_α, p_α, are the coordinates and momenta necessary to specify completely the classical state of molecule α. For the case of identical molecules the Hamiltonians, H_α, of (3) are all identical functions, the energies $\varepsilon_{m\alpha}$ will form the same energy spectrum for each of the identical molecules.

The Hamiltonian of the individual molecule is itself separable into one translational term,

$$H_{\text{tr}} = (1/2m)(p_x^2 + p_y^2 + p_z^2) \tag{2.1.13}$$

giving the kinetic energy due to the center of mass coordinates x, y, z, plus another term, H_i, of the internal coordinates and momenta relative to the center of mass. The internal Hamiltonian, H_i, is discussed in Chapter 3, and is itself usually approximately separable. The energy of the single molecules is given, then, by a sum of at least two terms, as in (1.5.4), one of them is ε_k due to the translational motion of the center of mass, and the other set is the term, ε_n, due to the internal degrees of freedom. The partition function Q_α of the individual molecule then also becomes a product, $Q_\alpha = Q_{tr}Q_i$, of which the translational term Q_{tr} has been evaluated in (1.5.18).

However, the quantum number $\boldsymbol{K}\dagger$ described by eq. (7) is not generally the same as the quantum number $\boldsymbol{K}$ describing the allowed states of (1.3.3). As a result one cannot identify ln Q_c with the logarithm of the sum over $\boldsymbol{K}\dagger$ given by (12). The reason is that the sums over the quantum numbers $\boldsymbol{m}_\alpha$ for the different identical atoms α are not completely independent. This is obvious for the Fermi–Dirac case, in which at most one molecule α can be in a single state $\boldsymbol{m}$. The lack of independence of the $\boldsymbol{m}_\alpha$'s is less obvious but also true for the Bose–Einstein case. The reason lies in the fact that eq. (5), stating that $\boldsymbol{K}\dagger$ is given by the set of numbers of the quantum states of the individual membered molecules, α, is erroneous for the allowed quantum members $\boldsymbol{K}$, of the system. The correct statement is that used in (1.5.5), namely that $\boldsymbol{K}$ is given by the numbers n_m, of molecules in the state $\boldsymbol{m}$. We discuss this in more detail in the next section.

2.2. Bose–Einstein, Fermi–Dirac and Boltzmann systems

The time independent Schrödinger equation (1.3.3) is a second order differential equation which must be solved subject to two different sets of conditions in order to find the states $\psi_{K\dagger}$ of the system. One of these sets of conditions is the set of boundary conditions, for instance that ψ vanish at the walls of the system. These conditions lead to a set $\boldsymbol{K}\dagger$ of quantum numbers of states $\psi_{K\dagger}$, with energies $E_{K\dagger}$.

The other conditions, the symmetry conditions, select from these functions, $\psi_{K\dagger}$, a limited number of linear combinations which are allowed. The symmetry conditions are that ψ must be antisymmetric (i.e. change sign only) if the coordinates of any two identical fundamental particles, neutrons, protons or electrons, namely the Fermions, are permuted. If the function ψ is written in the coordinates of molecules, then permutation of the coordinates of two identical molecules permutes as many pairs of identical Fermions as compose the molecule. If this number is odd the ψ must be antisymmetric in permutation of the molecular coordinates; the molecules are Fermi–Dirac in property. If the number of Fermions in the molecule is even, then ψ is symmetric in molecular permutations, and the molecules form a Bose–Einstein system.

The functions $\psi_{K\dagger}$ may have neither property. For instance a function of two molecules α and β which is a product, $\Psi_1 = \Psi_n(\boldsymbol{q}_\alpha)\,\Psi_m(\boldsymbol{q}_\beta)$, is neither symmetric nor antisymmetric if ψ_n and ψ_m are different functions. However, since the Hamiltonian is necessarily symmetric in permutation of $\boldsymbol{q}_\alpha$ and $\boldsymbol{q}_\beta$ if the molecules are identical, the function $\Psi_2 = \psi_m(\boldsymbol{q}_\alpha)\psi_n(\boldsymbol{q}_\beta)$ is also necessarily a solution to the Schrödinger equation with the same energy as Ψ_1. The linear combinations

$$\Psi_s = (1\sqrt{2})(\Psi_1 + \Psi_2) \tag{2.2.1}$$

$$\Psi_A = (1/\sqrt{2})(\Psi_1 - \Psi_2) \tag{2.2.2}$$

are also solutions to this energy, and are symmetric and antisymmetric, respectively. Only one of these, depending on the system, is allowed.

More generally, if there are N identical molecules, we form the $N!$ permutation operators, $\mathscr{P}_s$, $1 \leqslant s \leqslant N!$, which operate on the eigenfunctions by permuting the coordinates of the N identical molecules in the $N!$ different possible permutations. The operators include the one identical permutation which does nothing. For any single $\Psi_{K\dagger}$ the function defined by

$$\Psi_s = (1/\sqrt{A_s}) \sum_{s=1}^{s=N!} \mathscr{P}_s \Psi_{K\dagger}, \tag{2.2.3}$$

is necessarily symmetric and necessarily a solution of the Schrödinger equation. The function

$$\Psi_A = (1/\sqrt{A_A}) \sum_{s=1}^{s=N!} (-)^{p(s)} \mathscr{P}_s \Psi_{K\dagger}, \tag{2.2.4}$$

in which $p(s)$ is the parity of $\mathscr{P}_s$, namely zero if an even number of pairs are permuted and unity if an odd number, is antisymmetric, and also a solution to the same energy as $\Psi_{K\dagger}$.

If the permuted functions, $\mathscr{P}_s\Psi_{K\dagger}$, are different and mutually orthogonal, the normalization factor A is $N!$, which, however, does not concern us here. *If*, however, all the $N!$ permuted functions $\mathscr{P}_s\Psi_{K\dagger}$ are different and orthogonal they will have occurred as functions, $\Psi_{L\dagger}$, with a quantum number $\boldsymbol{L}\dagger$, different from $\boldsymbol{K}\dagger$, as defined by (2.1.6), since different numbered molecules exchange quantum numbers. Out of these $N!$ different functions we can construct only *one* totally symmetric function, or only *one* totally antisymmetric one. This means that we can, approximately at least, sum over all values of $\boldsymbol{K}\dagger$, and divide by $N!$, since all the $N!$ different Ψ's that enter the sums (3) or (4) have the same energy $\boldsymbol{E}$. In the general case, we divide by

$$N! = N_a!\,N_b \ldots, \tag{2.2.5}$$

and write

$$Q_c(V,T,N) = \sum_{K}^{(N,V)} e^{-E_K/kT} \cong (1/N!) \sum_{K\dagger}^{(N,V)} e^{-E_{K\dagger}/kT}, \qquad (2.2.6)$$

where the sum over $K\dagger$ includes all the states found as solutions of the Schrödinger equation without the limitation on selecting totally symmetric or antisymmetric functions in permutations of the identical molecules. The approximation of eq. (6) is known as the Boltzmann approximation, or frequently referred to as Boltzmann statistics.

If we use (6) and eq. (2.1.12) for the summation of $K\dagger$ in the case of the one component perfect gas, for which the α's are the individual identical molecules, and $Q_\alpha = Q_{\text{tr}}Q_i = V(2\pi mkT/h^2)^{3/2}Q_i$, from (1.15.18) one has

$$\ln Q_c(V,T,N) = N \ln V(2\pi mkT/h^2)^{3/2}Q_i - \ln N!. \qquad (2.2.7)$$

With the Stirling approximation that

$$\ln N! = N(\ln N - 1), \qquad (2.2.8)$$

and (1.4.12) for A one has

$$A(V,T,N) = NkT\{[\ln(N/V)(h^2/2\pi mkt(^{3/2}Q_i^{-1}]-1\}. \qquad (2.2.9)$$

Differentiation with respect to N at constant V and T gives the chemical potential μ as

$$\mu = (\partial A/\partial N)_{V,T} = kT \ln(N/V)(h^2/2\pi mkT)^{3/2}Q_i^{-1}. \qquad (2.2.10)$$

Since, from (1.5.20) $N/V = P/kT$, this is in agreement with (1.5.19) for μ.

The approximation of eq. (6) then leads to the same result as was obtained in section 1.5 by writing $\mp \ln(1 \mp e^{(\mu-\epsilon)/kT})$ as $e^{(\mu-\epsilon)/kT}$. This approximation is numerically very good for all real gases except helium at very low temperatures, and is even then quite good as long as it remains gaseous.

Let us examine in more detail the approximation of eq. (6). If the state functions, $\Psi_{K\dagger}$, are obtained as they are for a perfect gas it is easy to see just when the permuted functions $\mathscr{P}_s \Psi_{K\dagger}$ are really all different and orthogonal. In this case the $\Psi_{K\dagger}$ are products of the functions for the independent molecules,

$$\Psi_{K\dagger}(\mathbf{q}^{(N)}) = \underset{\alpha}{\pi} \psi_{m\alpha} \quad (\mathbf{q}_\alpha). \qquad (2.2.11)$$

The molecular state functions, $\psi_{m\alpha}$, for different quantum numbers $\mathbf{m}$ are all orthogonal. If all molecules α are in *different* quantum states $\mathbf{m}_\alpha$, the permuted function $\mathscr{P}_s\Psi_{K\dagger}$ changes the quantum states of the individual numbered molecules, and the states for different permutations are orthogonal, and are all different distinct states which are separately counted in describing $K\dagger$ by giving the quantum states of the numbered molecules, eq. (2.1.7).

If, however, two molecules α and β have $m_\alpha = m_\beta$ in (11), but all others are in different states, then two permutations $\mathscr{P}_s$ and $\mathscr{P}_{s'}$, differing only in the permutation of α and β, will always lead to the same state $\Psi_{K\dagger}$. Only $(1/2)N!$ different and orthogonal functions occur in the sums (3) and (4). In the antisymmetric sum (4) the two permutations $\mathscr{P}_s$ and $\mathscr{P}_{s'}$ occur with different signs, since they differ in a single pair permutation only, and their sum cancels to zero. Since for every permutation, s, there is a corresponding s', the complete sum is zero, which leads us to the well-known statement that for Fermi–Dirac systems no two identical molecules may have identical quantum states.

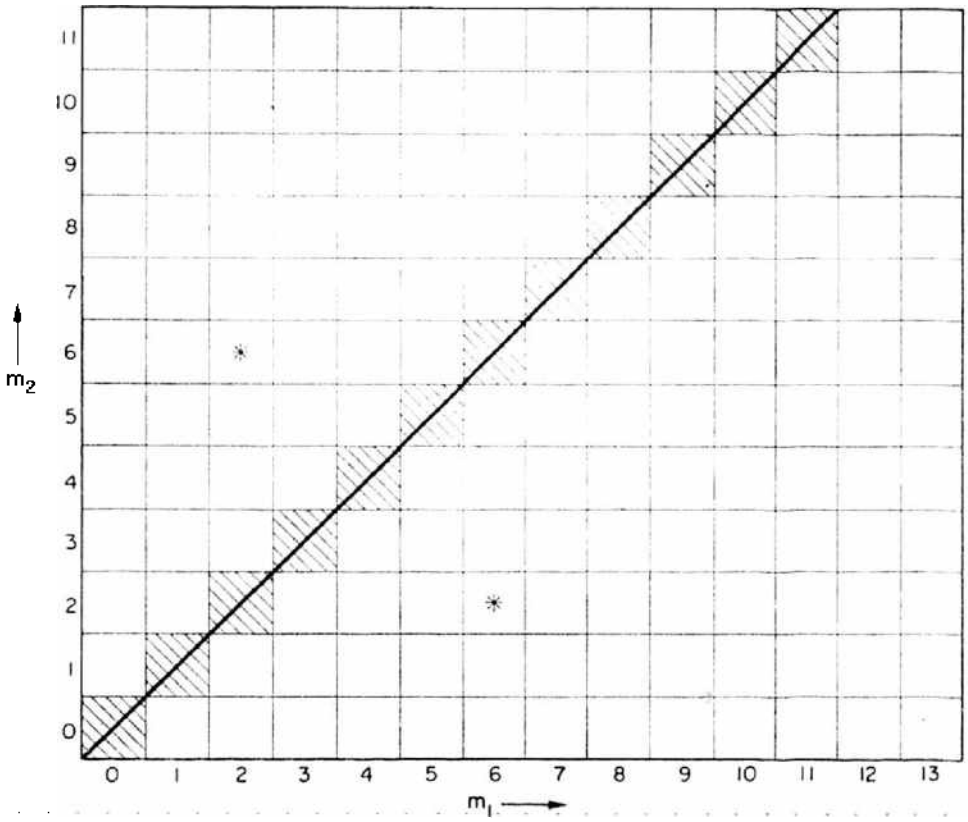

FIG. 2.1. Diagrammatic representation of quantum states for two identical molecules in one dimension. The shaded diagonal cells are unallowed to Fermions, allowed to Bosons, and counted half in Boltzmann counting.

In the Bose–Einstein system, however, the state $m_\alpha = m_\beta$ is allowed, but less than $N!$ different states $\Psi_{K\dagger}$ are required to construct it. The situation is most easily visualized for two identical independent molecules in the hypothetical one-dimensional case that we can describe m by a single number set, 0, 1, 2, 3, . . . (see Fig. 2.1). The $K\dagger$ value is given by m_1 and m_2, the state $K\dagger'$ of m_1', m_2', where $m_1' = m_2$, $m_2' = m_1$ is counted as different. Actually these two states, if $m_1 = m_2$ are combined by (1) or (2) to give a single symmetric *or* a single antisymmetric state as required. The true state is described, not by giving m_1 and m_2, but as in section 1.5, by saying that there is one molecule in m_1 and one in m_2, all other states m_ν have none. For these states the correct counting is obtained by summing over all m_1 and m_2 and dividing by $N! = 2! = 2$.

However, the diagonal states, $m_1 = m_2$ are unique. They must be omitted in the counting for Fermi–Dirac systems, but counted *without* division by two for Bose–Einstein systems. We always find that the Q_c computed by the Boltzmann approximation lies between the two cases. The Fermi–Dirac Q_c is less than the Boltzmann, the Bose–Einstein Q_c is greater. If the summation goes very far from the origin, that is if $\varepsilon_m << kT$ until m is very large, the error is small, since the diagonal points become a negligible fraction of all. If T is low, or ε_m increases rapidly with m, as it does from (1.5.3) for very small masses, then the error may become appreciable in using the Boltzmann approximation.

The correct counting, as we saw in section 1.5, is easily enough performed for perfect gases by using the grand canonical ensemble. Actually, eq. (1.5.11), although not quite as convenient as the Boltzmann approximation of eq. (1.5.12), can be used readily (Chapter 7). However, for systems in which the molecules interact, for which the states of the system are not simply describable as products of single molecular functions, more mathematical ingenuity is required to make corrections to the Boltzmann approximation. Fortunately the corrections are seldom required in chemical problems. For most molecules at ordinary temperatures the Boltzmann approximation is excellent. Two cases of importance in which the approximation is poor are those of liquid helium and electrons in metals, the latter even at very elevated temperatures.

In the next section we discuss the classical limit of these quantum mechanical equations. Each quantum state corresponds to a cell of volume h^Γ in momentum coordinate space. The states in which two or more molecules have the same quantum state correspond to the portions of the classical phase space in which these molecules are spatially close, and moving with equal or nearly equal momenta. One frequently describes the different systems qualitatively by the statement that Fermi–Dirac systems have a statistical repulsion between molecules in phase space, and Bose–Einstein systems have a statistical attraction.

2.3. The classical limit

In classical mechanics the mechanical state of a system is defined at any time, t, by giving the precise values of the Γ dimensional coordinates, $q^{(N)}$, and momenta, $p^{(N)}$. The future of the system is then precisely determined by the laws of mechanics. The space of coordinates and momenta is called the "phase space." The phase space of the whole system is referred to as the γ-space[†] (gas space), and has some 10^{20} or more degrees of freedom. The phase space of a single molecule, which has a more reasonable number of degrees of freedom is called the μ-space (molecular space). The classical state of the system is given by one point in λ-space, or for N molecules of the same species by N points in μ-space. If several molecular species are present in the system the μ-spaces of the different species usually have differing numbers of degrees of freedom, and a different μ-space must be used for each species. If, however, as is sometimes the case, one is describing only the positions and momenta of the centers of mass of the molecules, a single μ-space of three momenta and three coordinates may be used for all species, and the state of the system described by giving N_a points of type a, N_b of type b, . . ., etc.

The quantum mechanical state of a system is described by giving a quantum number L, which in turn defines some state function, Φ_L. In the general case the state function Φ_L is time dependent, and can be written as a linear combination of the stationary Ψ_K's, each multiplied by a (complex) coefficient, $a_{L,K}$, and the time dependent factor $\exp i(E_K/\hbar)t$. For purposes of the discussion of equilibrium systems it suffices to focus our attention on the stationary states alone.

The individual stationary states $\Psi_{K\dagger}$ can be chosen so as to correspond to a region of volume h^Γ in the γ-space. That is, the average values $<F_K>$, of all functions $F(p^{(N)}, q^{(N)})$, are such that they would correspond to a classical probability function in γ-space, more or less localized to some region of volume h^Γ. More precisely, the statement may be made as follows. Choose a volume $\Delta W >> h^\Gamma$ in the γ-space, such that the projected area, Δw_i, on any conjugate momentum coordinate plane, p_i, q_i, is greater than h, $\Delta w_i >> h$. The behavior of a classical ensemble with equal probability density in this volume element, and zero elsewhere, will be the same as that of the quantum mechanical ensemble having equal probabilities in each of $\Delta W/h^\Gamma$ different states, Φ_L. These states will not, in general, be the eigenstates of fixed energy, but selected wave packets corresponding to the classical region ΔW.

[†] The nomenclature was introduced by Ehrenfest.

This means that the sum over the states $K\dagger$ corresponds to the integral over γ-space,

$$\sum_{K\dagger} \cong h^{-\Gamma} \int\int \cdots \int dp^{(N)}\, dq^{(N)}. \tag{2.3.1}$$

This replacement becomes asymptotically exact if the summation is over a function which varies negligibly, in each degree of freedom, i, in regions of phase space of area h. That is we can replace

$$\sum_{K\dagger}^{(N,V)} e^{-E_{K\dagger}/kT} \cong h^{-\Gamma} \int\int_V \cdots \int \left[\exp - \frac{1}{kT} H_N(p^{(N)}, q^{(N)})\right] dp^{(N)} dq^{(N)}, \tag{2.3.2}$$

with H_N, the classical Hamiltonian Function for the set N of molecules in V. The replacement is numerically excellent if H_N is such that for every degree of freedom, i, in the regions of phase space which contribute appreciably to the integral, one can find areas,

$$\Delta w_i = \int\int dp_i\, dq_i >> h, \tag{2.3.3}$$

such that H_N varies negligibly compared with kT within the region.

Using the Boltzmann approximation, (2.2.6) one then has

$$Q_c(V,T,N) \cong \frac{1}{N!h^\Gamma} \int\int \cdots \int \left[\exp - \frac{1}{kT} H_N(p^{(N)}q^{(N)})\right] dp^{(N)} dq^{(N)} \tag{2.3.4}$$

as the classical approximation for the macroconical partition function.

The replacement of integration in eq. (1) for the summation refers to the summation over the states, $K\dagger$, selected without regard to the symmetry condition. Mathematical terminology requires that the different components of the momenta and coordinates, $p^{(N)}$, $q^{(N)}$, be numbered, so that if these components are those of individual molecules the individual molecules are assigned specific numbers, 1, 2, 3, . . ., N, even if they are identical. A point in γ-space then specifies the position in μ-space of each numbered molecule. There are $N!$ different positions, differing only in permutations of identical molecules, which refer to physically identical situations. One position in γ-space corresponds to putting N *numbered* points in μ-space, the $N!$ different, but physically identical points of γ-space correspond to specifying in μ-space the location of N_a unnumbered points of type a, N_b of type b, etc.

Since the γ-space is a truly continuous space the volume element corresponding to counting only the $N!$ mathematically different but physically identical points once, would be exactly corrected for by division by $N!$. However, the states $\boldsymbol{K}\dagger$ corresponds to "cells" of volume h^{Γ} in γ-space, and the $N!$ different states $\boldsymbol{K}\dagger$ that enter the sum (3) or (4) correspond to the $N!$ different cells of γ-space obtained by permuting identical molecules. Those on the diagonals, where two or more identical molecules have the same quantum numbers, now correspond to non-zero volume elements of volume h^{Γ}. Since the integration is introduced by (1) as an approximate method for counting cells the integration near the diagonals, where two or more molecules have the same p and q requires special handling. In the case of Fermi–Dirac systems it should be omitted, in Bose–Einstein systems it should be given extra weight.

Equation (4) is thus a valid approximation only if two independent conditions are satisfied. These are:

(A) The quantum states corresponding to every degree of freedom must have energy differences small compared with kT, as discussed after eq. (2), and,

(B) The density of points describing identical molecules in μ-space must be low when the volume element is measured by $\Pi_i h^{-1}\, dp_i\, dq_i$, so that the probability of finding two identical molecules in the same state is not appreciable.

Needless to say, when the Hamiltonian is separable, the replacement of integration for summation is valid for any degree of freedom which satisfies condition (A) above. We have already used this in computing Q_{tr} for the single molecules to find (1.5.18). The classical kinetic energy is $(1/2m)p^2$, with $p^2 = \boldsymbol{p}\cdot\boldsymbol{p}$, and the phase space is given by $\boldsymbol{r}, \boldsymbol{p}$. The volume element $d\boldsymbol{r}\, d\boldsymbol{p}$ is $d\boldsymbol{r}\, 4\pi p^2 dp$. The integrand, $\exp -(1/2mkT)p^2$, does not depend on $\boldsymbol{r}$, so that integration of $d\boldsymbol{r}$ over the volume V gives the factor V: integration over dp gives $(2\pi mkT)^{3/2}$; and division of h^3 for the three degrees of freedom gives (1.5.18).

2.4. Heat capacity and equipartition

We shall now proceed to derive a law more honoured in the breach than in the observance: the law that, in the classical limit, each degree of freedom contributes additively $\frac{1}{2}k$ to the heat capacity due to its kinetic energy, in addition to any contribution from the potential energy.

The coordinates and momenta, $q^{(N)}$, $p^{(N)}$, may always be chosen as the cartesian coordinates, $\boldsymbol{r}_i = x_i, y_i, z_i$, $\boldsymbol{p}i = p_{xi}, p_{yi}, p_{zi}$, for the individual atoms i, $1 \leqslant i \leqslant N$ that compose the system. We shall tacitly assume, as indeed is very frequently the case, that only one electronic state of sufficiently low energy is present, and that electronic coordinates may be forgotten. We then use the classical eq. (2.3.4), and note than the Hamiltonian, when cartesian coordinates are used, has the form

$$H_N(\boldsymbol{p}^{(N)}, \boldsymbol{q}^{(N)}) = \sum_{i=1}^{i=N} \frac{1}{2m} p_i^2 + U_N(q^{(N)}), \tag{2.4.1}$$

$$p_i^2 = p_{xi}^2 + p_{yi}^2 + p_{zi}^2, \tag{2.4.2}$$

where U is the potential energy, which is a function of the coordinates alone. The Hamiltonian is separable in the single atomic momenta, and the complete set of coordinates. The integration over the momenta is now trivial. Assign the factor h^{-3} for the three degrees of freedom to this integral to obtain,

$$h^{-3} \int_0^\infty 4\pi p^2 e^{-p^2/2m_ikT}\, dp = (2\pi m_i kT/h^2)^{3/2} = \lambda_i^{-3} \tag{2.4.3}$$

for each atom i of mass m_i. Define the configuration integral as

$$Q_\tau = \frac{1}{N!} \iint_V \cdots \int \left[\exp - \frac{U_N(\boldsymbol{q}^{(N)})}{kT}\right] dq^{(N)}. \tag{2.4.4}$$

With (1.4.12) for $A = -kT \ln Q_c$, and (2.3.4) for Q_c one has

$$Q_c = (1/N!)\, (\underset{a}{\pi}\lambda_a - 3N_a)Q_\tau, \tag{2.4.5}$$

$$A = \sum_a 3N_a kT \ln\lambda_a - kT \ln Q_\tau. \tag{2.4.5$'$}$$

Now

$$TS = -T\left(\frac{\partial A}{\partial T}\right)_{V,N} = -A + \frac{3}{2} NkT - kT \frac{\partial \ln Q_\tau}{\partial \ln \mathrm{T}}, \tag{2.4.6}$$

and

$$E = A + TS = kT\left[\frac{3}{2} N + (\partial \ln Q_\tau / \partial \ln T)\right]. \tag{2.4.7}$$

The next derivative with respect to T is

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{V,N} = \frac{3}{2} Nk + \frac{\partial}{\partial T}\frac{kT^2}{Q_\tau}\frac{\partial Q_\tau}{\partial T}$$

$$= k\frac{3}{2}N + k\left(\frac{1}{kTQ_\tau}\right)^2\left[Q_\tau\left(kT^2\frac{\partial}{\partial T}\right)^2 Q_\tau - \left(kT^2\frac{\partial}{\partial T}Q\right)^2\right]. \quad (2.4.8)$$

The heat capacity at constant volume, C_v, then has the contribution $(1/2)k$ due to the kinetic energy of each degree of freedom, and since each atom of the system contributes three degrees of freedom, the total contribution due to kinetic energy alone is $3\,Nk/2$.

We shall now proceed to show that the second term, due to the potential energy is necessarily positive.

From the expression (4) for Q_τ, one sees that

$$kT^2\frac{\partial}{\partial T}Q_\tau = \frac{1}{N!}\int\int_V \cdots \int U_N\left[\exp - \frac{U_N}{kT}\right] dq^{(N)}, \quad (2.4.9)$$

$$\left(kT^2\frac{\partial}{\partial T}\right)^2 Q_\tau = \frac{1}{N!}\int\int_V \cdots \int U_N^2\left[\exp - \frac{U_N}{kT}\right] dq^{(N)}. \quad (2.4.10)$$

Now write $[kT^2(\partial Q_\tau/\partial T)]^2$ as a double integral over coordinates $q^{(N)}$ and $q'^{(N)}$ with potential $U_N = U_N(q^{(N)})$ and $U_N' = U_N(q'^{(N)})$ respectively. Similarly write $Q_\tau[kT^2\partial/\partial T]^2 Q_\tau$ as a double integral over $q^{(N)}$ and $q'^{(N)}$, with the factor U_N^2 in the integrand. Clearly we could also have written this with U'^2_N instead, or with $\frac{1}{2}(U_N^2 + U'^2_N)$. Since

$$\tfrac{1}{2}[U_N^2 - 2U_N U'_N + U'^2_N] = \tfrac{1}{2}(U_N - U'_N)^2,$$

we have

$$Q_\tau(kT^2\partial/\partial T)^2 Q_\tau - (kT^2\partial Q_\tau/\partial T)^2$$

$$= \frac{1}{2}\left(\frac{1}{N!}\right)^2\int\int \cdots \int [U_N - U_N']^2\left[\exp - \frac{U_N + U_N'}{kT}\right] dq^{(N)}dq'^{(N)}, \quad (2.4.11)$$

which is necessarily positive, since it is the integral of a positive integrand.

The second term in (8) due to the potential energy has a simple physical interpretation, which has an exact analogue in the more general quantum mechanical case. The term $\exp - E_K/kT$ in Q_c, eq. (1.4.4) is proportional to the probability that the system of fixed N in V at T will be in the quantum state $\boldsymbol{K}$, and $\exp - H_N/kT$ in the classical case is proportional to the probability density that the system will have the coordinate-momenta values $\boldsymbol{q}^{(N)}$, $\boldsymbol{p}^{(N)}$. With the separable Hamiltonian of eq. (1) this probability density is a product of a term $\exp =$ (kinetic energy$/kT$) and the term $\exp - U_N/kT$, which means that the probability densities in coordinate and momentum space are independent. The classical probability density of a single cartesian momentum component, p_{xi}, is always proportional to the Gaussian, $\exp\text{-}(p_{xi}{}^2/2m_ikT)$, independently of the position in configuration space, or of other momenta values. The classical probability density in configuration space is proportional to $\exp - U_N/kT$ and the proportionality factor is $Q_\tau{}^{-1}$.

$$W(\boldsymbol{q}^{(N)}) = \frac{1}{N!Q_\tau} \exp -(U_N/kT), \tag{2.4.12}$$

for the probability density that the classical system of fixed N in V at T will have the coordinate value $\boldsymbol{q}^{(N)}$. This, then, is correctly normalized so that,

$$\int\int_V \cdots \int W(\boldsymbol{q}^{(N)})\, d\boldsymbol{q}^{(N)} = 1. \tag{2.4.13}$$

It is now clear from the definition of an average, using (13) and (9), that

$$<U> = -Q_\tau{}^{-1}\, kT^2(\partial Q/\partial T), \tag{2.4.14}$$

and from (10),

$$<U^2> = Q_\tau{}^{-1}\, (kT^2\partial/\partial T)^2 Q_\tau. \tag{2.4.15}$$

Using these with (8) one finds for the classical potential energy contribution, C_{vU}, to C_v,

$$C_{vU} = k(kT)^{-2}\,[<U^2> - <U>^2]. \tag{2.4.16}$$

For the more correct quantum mechanical case the contributions of the kinetic and potential energies are not necessarily separable. However, the total contribution to C_v can be written analogously to (16). One has, for the probability, $W(\boldsymbol{K})$, of the quantum state $\boldsymbol{K}$ in the closed system of N in V at T, the expression,

$$W(\boldsymbol{K}) = Q_c{}^{-1} \exp - E_K/kT. \tag{2.4.17}$$

Since, from (1.4.12) that $A = -kT \ln Q_c$, one has for the energy,

$$E = A + TS = A - T(\partial A/\partial T)_{v,N}$$
$$= T^2(\partial/\partial T)_{v,N}(A/T) = kT^2 Q_c^{-1}(\partial Q_c/\partial T). \tag{2.4.18}$$

With (17) for $W(\boldsymbol{K})$ and (1.4.4) for Q_c one sees that the thermodynamic energy of (18) is the average energy,

$$E = <E> = Q_c^{-1} \sum_{\boldsymbol{K}} E_{\boldsymbol{K}}\, W(\boldsymbol{K}), \tag{2.4.19}$$

as we also found in section 1.3 for the open system. Differentiating (18) again one has

$$C_v = (\partial E/\partial T)_{V,N} = k(kTQ_c)^{-2}[Q_c(kT^2\partial/\partial T)^2 Q_c - (kT^2 \partial Q_c/\partial T)^2] \tag{2.4.20}$$

analogously to the term C_{vu} in (8). Again one has, forming the average square energy, $<E^2>$ from (17)

$$<E^2> = Q_c^{-1} \sum_{\boldsymbol{K}} E_{\boldsymbol{K}}^2 \exp - E_{\boldsymbol{K}}/kT, \tag{2.4.21}$$

that

$$C_v = k(kT)^{-2}[<E^2> - <E>^2]. \tag{2.4.22}$$

2.5. Symmetry number

We have discussed the division of $N!$ as though no ambiguity could arise. We shall here raise a straw man, and demolish him. Should we divide by the factorial of the number set of atoms or of molecules?

Consider a system described by having a set $\boldsymbol{N} = N_a, N_b, \ldots$ of molecules of kind $a, b, \ldots$ etc. We may also describe the system by its number set $\boldsymbol{M} = M_\alpha, M_\beta, \ldots$ of atoms of kind α, β, etc. If each molecule of species a has $n_{a\alpha}$ atoms of kind α we have

$$M_\alpha = \Sigma N_a n_{a\alpha}. \tag{2.5.1}$$

Were we to use coordinates for the individual atoms it would be necessary to divide the sums over $\boldsymbol{K}$†, or the integrals over the phase space, by $\boldsymbol{M}! = \prod_\alpha M_\alpha!$ On the other hand, if our previous description of using molecular coordinates, and dividing by $\boldsymbol{N}! = \prod_\alpha N_a!$ is correct, the two must be consistent.

We first use the atomic coordinates, numbering the M_α atoms of kind α from 1 to M_α, those of kind β from $M_\alpha + 1$ to $M_\alpha + M_\beta$, etc., set up the Hamiltonian, and proceed to use (2.3.4) with division by $\boldsymbol{M}!$ We then observe that H is very large, and $\exp - H/kT$ negligible in value, except in the limited region of phase space for which the atoms are associated with molecules with $n_{a\alpha}$ atoms of kind α in the molecules of kind a, etc.

We therefore wish to introduce (necessarily numbered) molecular coordinates of the centers of mass of each molecule, and internal coordinates which specify the configuration of the atoms within the molecule. We must, however, in our original integral over the atomic coordinates, integrate over all possible ways of combining the numbered atoms in molecules. We can do this by multiplying the integral over the molecular coordinates by the number of ways we could assign the numbered atoms to numbered molecules and then divide by $N!$, since, in the integration over the molecular coordinates we include the phase space of all $N!$ permutations of the molecules, and these permutations have already been considered different in assigning the atoms to numbered *molecules.*

The number of ways that one can put M_α numbered objects in numbered piles, $n_{a\alpha}$ each in the first N_a piles, $n_{b\alpha}$ in the next N_b, . . ., is

$$M_\alpha!\prod_a [n_{a\alpha}!]^{-N_a}, \tag{A}$$

(see end of this section) and for all atomic species α the number is, with $\boldsymbol{n}_a! = n_{a\alpha}!\, n_{a\beta}! \ldots$,

$$\boldsymbol{M}!\prod_a (\boldsymbol{n}_a!)^{-N_a}. \tag{B}$$

We find, then,

$$\begin{aligned} Q_c(V,T,\boldsymbol{M}) &= \frac{1}{\boldsymbol{M}!h^\Gamma}\int\int\cdots\int\left[\exp - \frac{H_M}{kT}\right] d\boldsymbol{p}^{(M)}\, d\boldsymbol{q}^{(M)} \\ &= \prod_a(\boldsymbol{n}_a!)^{-N_a}\frac{1}{\boldsymbol{N}!h^\Gamma}\int\int\cdots\int\left[\exp - \frac{H_N}{kT}\right] d\boldsymbol{p}^{(N)}\, d\boldsymbol{q}^{(N)}, \end{aligned}$$

where the number of degrees of freedom, Γ, in the two phase space systems will be unchanged.

We must now, in the latter expression of integration over the molecular coordinates, integrate over all possible of the $\boldsymbol{n}_a! = \prod_\alpha n_{a\alpha}!$ permutations of the atoms of species α in each molecule of species a. Actually, the internal Hamiltonian, H_{ia}, of the molecules of species a, which is commonly used may not be such that this is accomplished correctly.

The detailed form of the Hamiltonians that are appropriate for various molecular species are discussed more fully in Chapter 3. One usually chooses internal coordinates that describe rotations of the molecule as a rigid frame, and some "vibrational" coordinates that specify displacements of the atoms from their equilibrium positions in this rigid equilibrium frame. The Hamiltonian is then chosen such that it correctly represents the true values for small displacements along these vibrational coordinates. It may not at all give the correct value for extremely large displacements that correspond to a new equilibrium configuration in which two or more identical atoms have been exchanged, but instead give extremely high values for such a displacement, so high that $\exp - H/kT$ is effectively zero. In this case the approximate Hamiltonian used leads to integrals that must be multiplied by the number of equilibrium minima not correctly represented, before division by $\boldsymbol{n}_a!$

For instance, suppose the molecule is ClHC=CBrH in the *cis* form, $n_{\mathrm{H}} = 2$, $n_{\mathrm{C}} = 2$, $n_{\mathrm{Cl}} = n_{\mathrm{Br}} = 1$ and $\boldsymbol{n}! = 4$. The usual internal H_i would be (approximately) correct for small vibrational displacements of the atoms, but would give utterly false high values for all except one of the four different configurations which would be written were numbers assigned to the two identical carbons and the two hydrogens. In this case we should multiply the integral by $\boldsymbol{n}! = 4$ for each molecule, precisely canceling the factor $(\boldsymbol{n}_a!)^{-N_a}$ in the expression for Q_c.

Now there are cases in which integration over the internal coordinates does include more than one of the $\boldsymbol{n}_a!$ configurations corresponding to permutations of the numbered identical atoms in the single molecule of kind a. In general they may include a number, σ, of them. The number σ is called the symmetry number. The correct expression for the partition function is then

$$Q_c(V,T,N) = \frac{1}{N!h^{\Gamma}\prod\limits_a \sigma_a^{N_a}} \iint_V \cdots \int e^{-H/kT}\, dp^{(N)}\, dq^{(N)}. \quad (2.5.2)$$

The symmetry number, σ_a, for molecules of species a, is then the number of permutations of the identical atoms composing the molecule of type a for which the internal Hamiltonian that is used does actually give the correct value. This number is less a characteristic of the molecule itself, than of the coordinate system and Hamiltonian which is used.

In treating diatomic elements the usual polar angles of rotation of the major axis, θ, ϕ include the two positions of the molecule that differ only in exchange of the two identical atoms, so that $\sigma = 2$. In the case of CH_4 with $n_c = 1$, $n_{\mathrm{H}} = 4$, $\boldsymbol{n}! = 1!\,4! = 24$, the integration over the internal coordinates used would naturally cover the complete rotational angles, which include $\sigma = 12$ different permutations of the four identical hydrogens, but not all 24 permutations. The topologically similar molecule with four different atoms, CHClBrI, has $2 = n!/\sigma = 24/12$ optical isomers. The other 12 permutations of the atoms are reached by rotation. However, were we to treat a crystal of methane at low temperatures, we might find it convenient to use a Hamiltonian which was only correct for small libations of the individual molecules within the crystal, and that did not give correct energies for angular displacements large enough to permute the hydrogen atoms. In this case σ for methane would shrink to unity.

The relation of the symmetry number to the rotational quantum states, and some more detailed examples are treated in section 3.8.

The somewhat awkward numbering of identical particles, and then correcting for the numbering, which was necessary in this section, is encountered frequently in statistical mechanics, particularly in the classical integrals. The counting of the number of ways that certain objects can be grouped together, or assigned to cells or boxes is usually most readily accomplished by assigning numbers to the objects and boxes, and then correcting for this assignment if necessary.

The argument that leads to the number of ways assigning the atoms to the molecules used here is typical of a number of combinatory problems. In detail it goes as follows.

We have M_α numbered objects of type α, and $N = \Sigma N_a$ numbered boxes; the first N_a take $n_{a\alpha}$ objects each, the second N_b take $n_{b\alpha}$ each, . . ., $\sum_a n_{a\alpha} N_a = M_\alpha$. We can arrange the M_α numbered objects in consecutive order in $M_\alpha!$ different ways, since we can choose any of the M_α for the first place, any of the $M_\alpha - 1$ remaining for second place, . ., any of the $M_\alpha - n$ remaining for the $(n + 1)$th place. To each such consecutive order we can uniquely assign an arrangement of the numbered objects in the numbered boxes, namely that with the first $n_{a\alpha}$ objects in the first box, the second $n_{a\alpha}$ in the second box, . . ., numbers $n_{a\alpha} N_a + 1$ to $N_a + n_{b\alpha}$ inclusive in the $(N_a + 1)$th box which is the first box of type b, . . ., etc. However, many of the original $M_\alpha!$ consecutive orders of the objects give the same assignment, namely all those $n_{a\alpha}!$ ones in which the first $n_{a\alpha}$ objects were permuted among themselves, and with each of these permutations all those in which the second $n_{a\alpha}$ are permuted, and with each of these . . ., etc. In all there are $\prod_a (n_{a\alpha}!)^{N_a}$ different consecutive orders of the numbered objects that correspond to the same assignment to the boxes. We therefore have expression (A) for the number of ways of putting the numbered objects in numbered boxes.

The same expression, with M_β and $n_{a\beta}$ replacing M_α and $n_{a\alpha}$ will be obtained for putting the objects of type β in the same boxes, and since every arrangement of the α objects can occur with every arrangement of the β objects the total number of ways is the product expression (B).

Now, however, if the boxes were really unnumbered we would have computed as above, but then divided by $\mathbf{N}!$ since every one of the $\mathbf{N}!$ different ways of arranging the boxes in order would have led to the same association of the numbered objects in boxes, N_a boxes having the set $n_a = n_{a\alpha} n_{a\beta} \ldots$ of $n_{a\alpha}$ objects of kind α, $n_{a\beta}$ of kind β, . . . etc.

We thus have

$$C\{N\} = \mathbf{M}!/\mathbf{N}\prod_a (\mathbf{n}_a!)^{N_a} \tag{2.5.3}$$

for the number of ways that we can arrange the set $\mathbf{M} = M_\alpha, M_\beta \ldots$, of numbered objects in the set $\mathbf{N} = N_a, N_b, \ldots$ of *unnumbered* boxes, with that $\mathbf{n}_a = n_{a\alpha}, n_{a\beta}, \ldots$ of objects in each box of type a.

2.6. Isotopes

The existence of isotopes having mole fraction $x_{\alpha i}$ for isotope i of chemical element α,

$$\sum_i x_{\alpha i} = 1 \tag{2.6.1}$$

introduces a positive entropy of mixing term,

$$\Delta S_\alpha^{(i)} = -kM_\alpha \sum_i x_{\alpha i} \ln x_{\alpha i} \tag{2.6.2}$$

per M_α atoms of element α in any system, but has no other effect *in the classical limit* on the properties of the system.

In the foregoing we discussed the effect of the identity of atoms of species α or molecules of species a. It is frequently, and indeed usually, inconvenient to label the species by both the elements *and* the isotopes present. Two atoms of differing isotopes of the same element are not identical, their interchange does not permute identical fundamental Fermions only, but in addition transfers one or more neutrons from one position to another. There is no requirement for the state function Ψ_K to be either symmetric or antisymmetric in the permutation of atoms consisting of differing isotopes. The number set M of numbers

M_α of atoms of the species α, then requires that α refer not only to the element, but also to the isotope, namely the set of $M_{\alpha i}$ for isotope i of element α.

Now, for the heavier elements—and for many purposes every element except hydrogen can be considered heavy—the mechanical properties of molecules differing only in their isotope composition are very closely similar. Two of the quantum states $K\dagger$ and $L\dagger$ differing only in permutations of differing isotopes of the same element will have very approximately the same energies $E_{K\dagger} = E_{L\dagger}$. Insofar as this is exact the value of Q_c is then larger by the factor, for every element α,

$$I_\alpha = M_\alpha!/\prod_i M_{\alpha i}! \tag{2.6.3}$$

than the value which would have been computed had one used the Boltzmann approximation using only the numbers M_α of atoms of element α, and not the numbers $M_{\alpha i}$ of the truly identical isotope α_i.

The logarithm of I_α is given by Stirling approximation as,

$$\begin{aligned}\ln I_\alpha &= M_\alpha(\ln M_\alpha - 1) - \sum_i M_{\alpha i}(\ln M_{\alpha i} - 1)\\ &= -M_\alpha \sum_i x_{\alpha i} \ln x_{\alpha i},\end{aligned} \tag{2.6.4}$$

where

$$x_{\alpha i} = M_{\alpha i}/M_\alpha. \tag{2.6.5}$$

Since $A = -kT \ln Q_c$ one finds the additive term,

$$\Delta A_\alpha^{(i)} = M_\alpha kT \sum_i x_{\alpha i} \ln x_{\alpha i}, \tag{2.6.6}$$

in A. Using $-(\partial \Delta A_{\alpha i}^{(i)}/\partial T)_{K,m} = \Delta S_\alpha^{(i)}$ one finds eq. (2) for the additive entropy term due to the isotopes.

This term does not affect the energy, $E = A + TS$, as it is also intuitively obvious that it should not. Within the accuracy to which the approximation that the quantum states of the system have the same energy for all isotopes is valid, the mole fraction of the isotopes will be the same in all molecular species. The term $\Delta S_\alpha^{(i)}$ will cancel in all entropy differences for two sides of a chemical reaction. In actual thermodynamic processes, even at low temperatures where the differences in the energies of the states for differing isotopes are of the order of kT, and appreciable fractionation of isotopes between differing molecular species would occur at equilibrium, there is seldom equilibrium attained between chemical reactions. As a result, unless special experimental situations are constructed to detect it, this isotopic mixing entropy does not enter into the usual thermodynamic equations. It is common to omit the term completely in most statistical mechanical computations of thermodynamic quantities.

2.7. Nuclear spin

Like the effect of isotopes, the existence of nuclear spins of quantum value $j_{\alpha i}$ for the isotope αi introduce an additive term

$$\Delta S_{\alpha i}^{(\mathrm{n.s.})} = M_{\alpha i}\, k \ln (2j_{\alpha i} + 1), \tag{2.7.1}$$

in the entropy for $M_{\alpha i}$ nuclides of species αi, and this term is usually neglected in computations of the entropy.

The nucleus αi which has the total spin quantum number $j_{\alpha i}$, integer or half integer, has a quantum number m describing the orientation of the spin which can take all values differing by integers between $-j_{\alpha i}$ and $+j_{\alpha i}$ inclusive. There are $2j_{\alpha i} + 1$ such values, as can be seen by simple counting. The only *direct* effect on the energy of the differing nuclear spin orientations is the scalar product of the very small nuclear magnetic moment with the magnetic field at the nucleus, which may be due to an external applied field, or to a field produced by the electrons. The value of these energies is always extremely small compared with kT for all T's above even as low a value as 0·1°K.

The effect of the nuclear spin is therefore simply to multiply the total number of quantum states by the factor $(2j_{\alpha i} + 1)$ for each nuclide αi, and, as we have seen in discussing the isotope mixing effect this adds an entropy term of the magnitude of eq. (1). Just as the isotope mixing entropy this term is usually undetectable in ordinary thermodynamic processes, and hence omitted in most computations.

However, there is indirect effect of the nuclear spin on the energies, which, in some cases, notably in H_2 at low temperatures is very appreciable. This arises from the effect of the nuclear spin state on the symmetry of the eigenfunction, and hence on the rotational states which are allowed. It is discussed in section 3.5.

2.8. Coordinate transformations

It would take us far too far into the details of classical mechanics to discuss fully the use of coordinate transformations and their conjugate momenta. However, a brief review, emphasizing the points of special interest in statistical mechanics, follows.

Except for particle spins, which are always quantized to have half integer or whole integer angular momenta in units $h/2\pi$, one may always use cartesian coordinates, x_i, y_i, z_i, of the individual mass points, i, in the problems which concern us. The kinetic energy, K, then has the form,

$$K = \sum_i \tfrac{1}{2} m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2), \tag{2.8.1}$$

with $\dot{x}_i$, $\dot{y}_i$, $\dot{z}_i$ the velocities. The total number of degrees of freedom, Γ, is then three times the number of particles introduced (plus the number of rather special particle spin coordinates which may be present).

Except when magnetic forces are considered, which is relatively rare, there is a potential, U, which depends on the coordinates alone, and the total energy is $K + U$.

The momentum, p_{xi}, conjugate to x_i is

$$p_{xi} = \partial K/\partial \dot{x}_i = m_i \dot{x}_i, \quad \dot{x}_i = p_{xi}/m_i. \tag{2.8.2}$$

Using (2) to eliminate the $\dot{x}_i$'s, the total energy can be expressed in terms of the momenta and coordinates, and this function is the Hamiltonian,

$$H(p^{(N)}, q^{(N)}) = \sum_i \frac{1}{2m_i} (p_{xi}^2 + p_{yi}^2 + p_{zi}^2) + U(x_1, \ldots, z_N). \tag{2.8.3}$$

Not infrequently the potential U is a complicated function of the cartesian coordinates, but some other set q_α, $1 \leqslant \alpha \leqslant \Gamma$, of coordinates

$$q_\alpha = q_\alpha(x_1, \ldots, z_N), \quad x_i = x_i(q_1, \ldots q_\alpha, \ldots, q_\Gamma),$$
$$\dot{x}_i = \sum_\alpha (\partial x_i/\partial q_\alpha)\, \dot{q}_\alpha, \tag{2.8.4}$$

can be found, such that U is greatly simplified if written as a function of these, $U = U(q_1, \ldots, q_N)$. Using (4) the kinetic energy, K, of eq. (1) can then be written as a function of the new q_α's and $\dot{q}_\alpha$'s. The momentum p_α, conjugate to the coordinate q_α is now defined by the relation,

$$p_\alpha = \partial K/\partial \dot{q}_\alpha. \tag{2.8.5}$$

The resulting Γ equations, (5), must then be solved to find the $\dot{q}_\alpha$'s as functions of q_α's and p_α's, and the result used in the expression for the kinetic energy to write this as a function $K(p_1, \ldots, p_\alpha, \ldots, p_\Gamma, q_1, \ldots, q_\alpha, \ldots, q_\Gamma)$ of the momenta p_α and coordinates q_α. This kinetic energy plus U is now the Hamiltonian in the new coordinates and their conjugated momenta.

There are, however, several points that interest us particularly. Firstly, the coordinates q_α need not have the dimensions of length, for instance dimensionless angular coordinates are often used. However, from (5) the dimension of the product, $p_\alpha \dot{q}_\alpha$ is that of an energy, so that the dimension of any coordinate multiplied by its conjugate momentum is that of energy × time, which is the dimension of h,

$$[p_\alpha q_\alpha] = \text{energy} \times \text{time} = [h]. \tag{2.8.6}$$

Thus the classical integral expression, (2.3.4), for Q_c is always dimensionless in any coordinate system, as it must be.

The second point of importance is that the Jacobian of the transformation of the volume element in phase space is always unity, that is

$$\prod_{i=1}^{i=N} dx_i dy_i dz_i dp_{xi} dp_{yi} dp_{zi} = \prod_{\alpha=1}^{\alpha=3N} dq_\alpha dp_\alpha. \tag{2.8.7}$$

This again must, of course, be the case, if eq. (2.3.4) is to be valid in any coordinate-momentum system.

However, a warning is needed here. The Jacobian, $J(x_i, q_\alpha)$, of the transformation of the coordinate volume elements,

$$\prod_{i=1}^{i=N} dx_i dy_i dz_i = J(x_i, q_i) \prod_{\alpha=1}^{\alpha=3N} dq_\alpha \tag{2.8.8}$$

is usually *not* unity. Since the kinetic energy in cartesian space is so simple, and the integration over the classical cartesian momenta always leads to the simple factor λ_i^{-3} of eq. (2.4.3), one may integrate first over the momenta in cartesian coordinates. It may then be desired to make a transformation of coordinates to simplify the configuration integral of (2.4.4). In this case, of course, the Jacobian $J(x_i, q_\alpha)$ of (8) must be used in the integration.

CHAPTER 3

PERFECT GASES AND THE INTERNAL PARTITION FUNCTION

3.1. Monatomic molecules

Before proceeding further with more general equations we shall show how the results derived in section 1.5 can actually be used to obtain useful information for perfect gases. The equation for the chemical potential,

$$\mu = kT \ln (P/kT)\, \lambda^3\, Q_i^{-1}, \tag{3.1.1}$$

eq. (1.5.19), is sufficient to determine all thermodynamic properties if the internal partition function (1.5.14)

$$Q_i = \sum_n \exp - \frac{\varepsilon_N}{kT} \tag{3.1.2}$$

is evaluated. The sum of (2) runs over all internal states n of the molecule, with energy ε_n.

For monatomic molecules the only excited states are electronic, and except for the hydrogen atom, their energies are not given by any simple analytic expression. However, the energies of the excited states above the ground state are of order of one or more electron volts, and ε/k is 11,606°K at one electron volt. Unless the problem is one involving temperatures of the order of 10^4 deg only very few, and usually only the ground level, contribute significantly to the sum (2) for Q_i. If the spectroscopic ground level has total angular momentum quantum number $J = 0$, as it does for the noble gases, for which the lowest level is ${}^1S_0$, this ground level is a single quantum state, and $Q_i = 1$. For the alkali metals the ground level is ${}^2S_{1/2}$, $J = 1/2$, the number, $2J + 1$, of states in the level is 2, and $Q_i = 2$ up to several thousand degrees. There are low lying excited states in a few elements. For instance in atomic fluorine, the ground level, ${}^2P_{3/2}$ with $J = 3/2$, of $2J + 1 = 4$ states, has associated with it a doublet level ${}^2P_{1/2}$, $2J + 1 = 2$ states at a low energy of 407 cm^{-1}, $\Delta\varepsilon/k = 585$°K, and $Q_i = 4 + 2 \exp - (\Delta\varepsilon/kT)$ with $\Delta\varepsilon$ the rather low excitation energy of the doublet.

In general, then, for monatomic molecules Q_i is rather simply computed as a sum of a very few terms, and is usually, at low temperatures, a small integer independent of T.

It is, however, slightly disturbing to find that eq. (2) for Q_i appears to diverge if the more usual naive rules are applied to the quantum mechanical computation of the energy levels. The reason is that the condition of requiring the electron function ψ to vanish only at $r = \infty$, results in an infinite number of bound states below the ionization potential of the atom. Even if these states have an energy $\varepsilon >> kT$, so that $e^{-\varepsilon/kT}$ is indeed very small for each state, the product of infinity times a finite number, however small, is infinite. The origin of difficulty is easily seen by examining the average distance $<r>$ of these states from the nucleus. If $<r>$ is much larger than the average distance, in the gas, between molecules, then the state is essentially fictitious, and should somehow be grouped with the free electron states of ionized electrons. These are then handled as a separate chemical component. The electron states are, in turn, perturbed by their interactions with the ions present, and the system may be handled as an imperfect gas, Chapter 4. At reasonable temperatures and densities one generally finds that a simple estimate, assuming the electrons and ions to be perfect gases, already shows that ionization would be negligible. In other words, were the atomic states assigned the correct boundary condition of vanishing at the finite walls of the system, and a proper counting made, the contribution of those near ionization as well as those above ionization normally vanish. At higher temperatures, where ionization is appreciable, the situation becomes more complicated if sufficient precision is required to correctly include the imperfect gas corrections between electrons and ions. Essentially, then, one has a Debye Imperfect Gas correction, section 4.7, and counts the discrete excited atomic states up to those having the radius of the inverse Debye κ.

3.2. Diatomic molecules

Stable diatomic molecules usually have a lowest electronic configuration which is singlet with all electrons paired, and excited electronic states have, almost always, such high energies that they are excited only at temperatures for which the molecules themselves are dissociated. With very few exceptions, then, the internal states are due only to rotation and vibration of the nucleii, and with quite satisfactory precision the Q_i can be readily computed.

The few exceptions to the statement that the lowest state is singlet include 0_2, which has two electrons with parallel spins, and is therefore a triplet state. The Q_i for 0_2 has then an extra factor 3 beyond that discussed below. If, as in 0_2, the electronic state has an angular momentum, the rotational spectrum is somewhat more complicated than that discussed below, but the classical limit, which is actually used, is the same. One other exception, which has a degenerate lowest electrical level is NO, with an odd number of electrons, for which the lowest level is $^2\Pi$, consisting of 2 states, with another low-lying excited $^2\Pi$ at an energy $\varepsilon/k = 174°\text{K}$. In cases such as this the parameters entering the rotational and vibrational states are generally different for the two electronic levels, and the total Q_i must be treated as a sum of that computed for the lowest electronic state plus $e^{-\varepsilon/kT}$ times the Q_i for rotation and vibration with the parameters of the excited state. Such complications occur, but are rare.

Neglecting the possible electronic complications, the coordinates necessary to describe the configuration of the diatomic molecule are six in number, and are constructed from the cartesian coordinates, x_i, y_i, z_i, x_j, y_j, z_j of atoms i and j respectively as X, Y, Z, for the center of mass, and three internal coordinates, of which one is a length ξ, and the other two are the polar angles of the axis of the molecule, θ and ϕ. They are, with $m = m_i + m_j$, the total mass,

$$\begin{aligned} X &= m^{-1}(m_i x_i + m_j x_j), \\ Y &= m^{-1}(m_i y_i + m_j y_j), \\ Z &= m^{-1}(m_i z_i + m_j z_j), \\ \xi &= r - r_0,\ r = [(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{1/2}, \\ \cos\theta &= (z_i - z_j)/r, \\ \sin\phi &= (x_i - x_j)/r\sin\theta. \end{aligned} \tag{3.2.1}$$

The three center of mass coordinates and momenta give the factor $Q_{\text{tr}} = V\lambda^{-3}$, eqs. (1.5.18) and (3.1.3), to the molecular partition function. The three internal degrees of freedom, ξ, θ, ϕ, p_ξ, p_θ, p_ϕ account for Q_i.

The potential energy depends only on the distance between the two atoms $r = r_0 + \xi$ and r_0 is chosen at that distance for which $u(r)$ is a minimum,

$$(du/dr)_{r=r_0} = 0. \tag{3.2.2}$$

Introduce the moment of inertia, I, as

$$I = I(\xi) = \mu r^2 = \mu(r_0 + \xi)^2; \quad \mu = m_i m_j/(m_i + m_j). \tag{3.2.3}$$

The classical Hamiltonian is

$$H = \frac{1}{2\mu} p_\xi^2 + u(\xi) + \frac{1}{2I}\left(p_\theta^2 + \frac{1}{\sin^2\theta} p_\phi^2\right) \tag{3.2.4}$$

Now, since I is a function of ξ this is not strictly separable. One assumes that the vibrations along ξ are of such small amplitude that it is adequate to replace I by the value I_0, at $\xi = 0$, $I_0 = \mu r_0^2$. In this case the Hamiltonian is separated into a term depending on ξ and p_ξ only, and a rotational term depending on the angles and their momenta. In addition one assumes that it is adequate to approximate $u(\xi)$ above its minimum value by

$$u(\xi) = (1/2)\kappa\xi^2, \tag{3.2.5}$$

with κ a force constant.

The Hamiltonian is now separable,

$$H = H_v(p_\xi, \xi) + H_{\text{rot}}\,(p_\theta, p_\phi, \theta, \phi), \tag{3.2.6}$$

$$H_v = (1/2\mu)p_\xi^2 + (\kappa/2)\xi^2, \tag{3.2.7}$$

$$H_{\text{rot}} = (1/2I_0)\,[p_\theta^2 + (1/\sin^2\theta)p_\phi^2]. \tag{3.2.8}$$

Since the Hamiltonian has been separated one writes

$$Q_i = Q_v Q_{\text{rot}}, \tag{3.2.9}$$

the logarithm is a sum, and there are two additive contributions to the thermodynamic functions, one due to the oscillation along ξ, and the other due to rotation.

3.3. The oscillator partition function

The Hamiltonian H_v of (3.2.7) is that of an harmonic oscillator of mass μ and force constant κ. The classical solution is sinusoidal vibration along ξ with frequency,

$$\nu = (1/2\pi)\sqrt{(\kappa/\mu)}, \tag{3.3.1}$$

namely

$$\xi = a \sin\,(2\pi\nu t + \Delta) \tag{3.3.2}$$

with a the amplitude. The energy depends on a and is $E(a) = (1/2)\kappa a^2$. The quantum mechanical solutions are described by a quantum number n, and have discrete energies ε_n. The energy of the lowest state, $n = 0$, lies above the minimum, u_0, of the potential energy

$$u_0 = u(\xi = 0) \tag{3.3.3}$$

by an amount $\Delta\varepsilon = (1/2)h\nu$. For most chemical purposes it is convenient to choose this state as the zero of energy, so that the energy of the perfect gas at $T = 0°\text{K}$ is zero. For this choice, $\varepsilon_0 = 0$, the nth state has energy,

$$\varepsilon_n = nh\nu, \quad (\varepsilon_0 = 0). \tag{3.3.4}$$

If, however, the chemical equilibrium between molecules of differing isotopic composition is being explicitly considered the above choice is awkward, and $u_0 = 0$ is a preferable zero. This arises because, for differing isotopic composition the u_0's are very closely identical in value, compared with free atom energies. Since the frequencies ν involve the reduced mass, μ, through (3.2.9) the energies ε_0 may differ appreciably compared with the dissociated atoms, or the lowest states of other chemical compounds.

Using the choice of eq. (4), $\varepsilon_0 = 0$, one has for the partition function

$$Q_v = \sum_{n=0} e^{-nh\nu/kT} = (1 - e^{-h\nu/kT})^{-1}, \tag{3.3.4'}$$

and for the additive contribution, μ_v, to the chemical potential μ, due to this degree of freedom

$$\mu_v = -kT \ln Q_v = kT \ln [1 - e^{-h\nu/kT}]. \tag{3.3.5}$$

At high temperatures for which $h\nu << kT$ one may develop,

$$e^{-x} = 1 - x + (x^2/2) - (x^3/6) + \dots$$

$$\ln(1 - e^{-x}) = \ln[x - (x^2/2) + (x^3/6)^{-}\dots] = \ln x + \ln[1 - (x/2) + (x^2/6) - \dots]$$

$$= \ln x - (1/2)x + (1/24)x^2 + \dots.$$

Using this in (5) one has, for high temperatures

$$\mu_v = (1/2)h\nu + kT[\ln(h\nu/kT) + (1/24)(h\nu/kT)^2 + \dots]. \tag{3.3.6}$$

Were the energy zero chosen as $u_0 = 0$, which is $(1/2)h\nu$ lower, the value would be

$$\mu_v(u_0 = 0) \simeq kT[\ln(h\nu/kT) + (1/24)(h\nu/kT)^2 + \dots]. \tag{3.3.7}$$

The classical partition function, with $u_0 = 0$ as the zero of energy, is

$$Q_v^{(\text{Cl})} = h^{-1}\iint_{-\infty}^{+\infty} \left\{\exp - \frac{1}{2kT}[\kappa q^2 + \mu^{-1}p^2\right\} dq\, dp$$

$$= (kT/h)(2\pi\sqrt{\mu/\kappa}) = kT/h\nu \tag{3.3.8}$$

from (1) for ν. The term $kT \ln(h\nu/kT)$ in (7) is then the classical limit, which is always that found at high temperatures.

The frequency ν is never actually obtained from the force constant, κ, through (1), as would be theoretically possible if the energy of the electronic state for all nuclear distances near r_0 were computed, and this energy used to compute $u(r)$ for the nuclei. Instead the observed spectroscopic values of the vibration frequency, ν, are used. The numerical values are high, and the classical approximation (6) is seldom useful for diatomic molecules. Indeed, for H_2 one has $h\nu/k = 5958°K$, for N_2, $h\nu/k = 3336.6°K$, and for O_2 $h\nu/k = 2228.0°K$. At temperatures below 700°K the approximation $Q_v = 1$ is adequate for these gases. Only for the heavier elements, I_2 with $h\nu/k = 305.1°K$, Br_2 with $h\nu/k = 460.4°K$, or the very weakly bound molecules such as Na_2 with $h\nu/k = 227.4°$, does the vibrational contribution become important at room temperature.

3.4. The linear rotator function

The classical solution to the Hamiltonian (3.2.8) of the rigid dumbbell rotator is any rotation of total angular momentum vector, $\mathbf{p}$, with any direction in space. The energy is $\varepsilon = (1/2\,I_0)|\mathbf{p}|^2$. The quantum states are described by two integer quantum numbers, j and m. The square angular momentum is $|\mathbf{p}|^2 = (h/2\pi)^2 j(j+1)$, and the energy is

$$\varepsilon(j,m) = (h^2/8\pi^2 I_0)\, j(j+1), \tag{3.4.1}$$

which is independent of the quantum number m. The quantum number m defines the component, p_z, of $\mathbf{p}$ along a unique axis z, as having the value $p_z = m(h/2\pi)$, and m may take any of the $2j+1$ integer values,

$$-j < m < +j. \tag{3.4.2}$$

The rotator partition function is then

$$Q_{\text{rot}} = \sum_{j=0}^{\infty} \sum_{m=-j}^{m=+j} e^{-\varepsilon(j,m)/kT} = \sum_{j=0}^{\infty} (2j+1)\, e^{-\xi j(j+1)}, \tag{3.4.3}$$

where

$$\xi = (h^2/8\pi^2 I_0 kT) = B_0 hc/kT. \tag{3.4.4}$$

The quantity

$$B_0 = h/8\pi^2 I_0 c \tag{3.4.5}$$

of dimension length^{-1} is that most frequently listed in tables of the constants for diatomic molecules. Its value is such that $B_0hc/k =$ 2.847°K for N_2, 2.059°K for O_2, 0.05340 for I_2 but has the large value 84.971°K for H_2. Except for H_2 the value of ξ is very much less than unity at temperatures for which the molecules are normally gaseous.

There is no analytical expression for the sum of eq. (3). If, however, $\xi << 1$, the terms of low j values lie close together in value, and one may approximate,

$$Q_{\text{rot}} = \sum_{j=0} (2j+1)e^{-\xi j(j+1)} = \int_0^\infty (2j+1)e^{-\xi j(j+1)}\,\mathrm{d}j$$
$$= \int_0^\infty e^{-\xi j(j+1)}\,d[j(j+1)] = \xi^{-1}, \tag{3.4.6}$$

which leads to

$$\mu_{\text{rot}} = -kT\ln Q_{\text{rot}} = kT\ln\xi. \tag{3.4.7}$$

The contribution to the thermodynamic energy per molecule is $\bar{\varepsilon}_{\text{rot}} =$ $\mu - Td\mu/dT$, and since ξ is proportional to T^{-1} one has $\bar{\varepsilon}_{\text{rot}} = kT$, in agreement with the classical requirement for two kinetic energy degrees of freedom, section 2.4.

By use of the Euler–Maclaurin summation formula one finds the somewhat better approximation,

$$Q_{\text{rot}} = \xi^{-1}[1 + (1/3)\xi + (1/15)\xi^2 + \ldots], \tag{3.4.8}$$

which can be used when ξ is small but not negligible compared with unity. For H_2 and D_2 at low temperatures, but practically for no other molecule, Q_{rot} must be evaluated by numerical summation of (3) directly, a chore which is not too tedious, since even at the low value $\xi = 0.1$, the tenth term, $j = 9$, has the very small value $19e^{-9}$.

As one expects the value of eq. (6) obtained by integration over the quantum number j, as an approximation to summation, is the same as the $Q_{\text{rot}}{}^{(\text{Cl})}$ obtained by integration over phase space. One has

$$Q_{\text{rot}}{}^{(\text{Cl})} = h^{-2}\int_0^\pi d\theta\int_0^{2\pi} d\phi\int\int_{-\infty}^\infty dp_\phi\,dp_\theta \exp -\frac{1}{2I_0kT}\left[p_\theta{}^2 + \frac{p_\phi{}^2}{\sin^2\theta}\right]$$
$$= h^{-2}(4\pi^2 I_0kT)\int_0^\pi \sin\theta d\theta = \xi^{-1}. \tag{3.4.9}$$

As discussed in (2.5), the classical integral over the internal coordinates of any molecule having two or more identical atoms must be divided by the symmetry number σ which depends on the nature of the internal coordinates used. In the case of the diatomic element of identical atoms the angular coordinates θ, ϕ, cover complete rotation of the molecule, which includes then, the positions in which the two identical atoms have merely changed places. For this case $\sigma = 2$, and the classical $Q_{\text{rot}}{}^{(\text{Cl})}$ is half of eq. (9).

$$Q_{\text{rot}}{}^{(\text{Cl})} \text{ (diatomic element)} = (1/2\ \xi). \qquad (3.4.10)$$

The quantum mechanical state function of the diatomic element of identical isotopes must be completely symmetrical or antisymmetrical in exchange of the nucleii, depending on whether the atomic weight of the isotope is even or odd, respectively. Since $m_i \equiv m_j$, the coordinates X, Y, Z, and ξ, of (3.2.1) are seen to be unchanged by exchange of the positions of i and j. The translational and vibrational eigenfunctions are then unchanged, namely symmetrical in exchange. However, $\cos\theta$ changes sign, $\theta \rightarrow \pi - \theta$, and $\phi \rightarrow \pi + \phi$, in exchange of the two nucleii. The rotational eigenfunctions are such that if j is odd they change sign, whereas if j is even they remain unchanged by this transformation, that is, the even-j functions are symmetric, and the odd-j functions antisymmetric in the nuclear exchange.

Thus a diatomic molecule composed of identical isotopes of one element may take only states of even, *or* only states of odd j-values. The sum of (6) for Q_{rot} is limited to half of the j-values in either case. Insofar as the sum can be approximated by the integral method of (6) it has half the value given by (6), namely that of eq. (10). The case of H_2, for which the classical limit is inapplicable below room temperature, is discussed in more detail in the next section.

3.5. Molecular hydrogen

The molecule H_2, because of its low mass, and relatively small value of r_0, has a much lower moment of inertia, I_0, than any other molecule, and a high value of B_0. The energy, $\varepsilon_1 = 2B_0hc$, of the first excited rotational level, $j = 1$, has the value $\varepsilon_1/k = 169.9$°K, so that even considerably above its low boiling point of 20°K no rotational levels are appreciably excited at equilibrium. However, there is an indirect effect of nuclear spin on the energy, which results in a curious complication. This type of complication is present, in principle at least, in all symmetrical molecules of elements having non-zero nuclear spin, at sufficiently low temperatures. However, in all cases other than H_2 (and D_2) the effect is negligible unless the temperature is so low that the molecules condense, in which case it is more difficult to disentangle. We discuss H_2 at some length, not only for its own interest, but to illustrate the general phenomenon.

As discussed in the last section, the translational and vibrational coordinates are unaffected by nuclear exchange, and the rotational states are symmetric if j is even, and antisymmetric if j is odd. The hydrogen nucleus has spin 1/2, and two states differing in orientation of the spin exist. If these states are called α and β, the spin state of the molecule of nucleus i and j may have any of the four states,

$$\left.\begin{aligned} \alpha(i)\alpha(j) &= \psi_{1,1}, \\ (1/\sqrt{2})[\alpha(i)\beta(j)+\beta(i)\alpha(j)] &= \psi_{1,0}, \\ \beta(i)\beta(j) &= \psi_{1,-1}, \\ (1/\sqrt{2})[\alpha(i)\beta(j)-\beta(i)\alpha(j)] &= \psi_{0,0}. \end{aligned}\right\} \qquad (3.5.1)$$

The first three of these states have total nuclear spin unity, but projection 1, 0, and -1, respectively, along the axis for which $\alpha = +1/2$ and $\beta = -1/2$. The last state has antiparallel spins, net spin zero, and of course zero projection. The three states with parallel spins are called ortho-hydrogen, the one state of antiparallel spins is called para-hydrogen.

The three spin states of ortho-hydrogen are symmetric in exchange of i and j. Since the proton nucleus of H is a Fermion, the complete state of the molecule must be antisymmetric in this exchange, and hence only antisymmetric rotational states of odd j values are allowed to ortho-hydrogen. Only even j values, with symmetric rotational states, are allowed for para-hydrogen, since the nuclear state is already antisymmetric. For two consecutive j values, then the even j occurs with one nuclear spin state, the odd j with three. At high temperatures the Q_{rot} will be $2\xi^{-1}$ instead of ξ^{-1}, since on the average twice as many quantum states, including the nuclear spin states, are present as we assumed in deriving (3.4.6). This factor two is in agreement with our classical limit rules; the internal partition function is multiplied by two for each atom if the spin states are counted, and divided by the symmetry number, $\sigma = 2$. If, then, as is frequently done, we wish to subtract the nuclear spin entropy we divide by 4 and have only the symmetry number term left.

However, at low temperatures, and at equilibrium the molecules will all be in the lowest rotational states of $j = 0$, and only the single para-hydrogen nuclear state $\psi_{0,0}$ is permitted. The nuclear spin entropy is "frozen out".

In section 2.7 we discussed nuclear spin effects, and stated that the direct energy difference of different orientations was negligible. We have here, however, an indirect effect of the nuclear spin orientation on the energy which involves a term $\varepsilon/k = 169.9°\text{K}$; namely that different spin orientations occur with neighboring rotational levels of the whole molecule. In general, then, the nuclear spin entropy can be affected, in symmetric molecules, at temperatures where degrees of freedom of rotation cease to be classical. The nuclear spin entropy of H_2 takes its full value at complete equilibrium only well above the boiling point.

There is, however, an additional complication. Transitions between the symmetric and antisymmetric spin states are not produced by static electric fields, but only by inhomogeneous magnetic fields which act differently on the magnetic moments of the two nucleii in one molecule. Collisions with ordinary diamagnetic molecules have practically zero effect in producing these transitions. Only paramagnetic impurities, such as O_2, or unsaturated surface atoms on a catalyst such as charcoal, cause reasonably rapid transitions. In the absence of such special catalysts the half time for conversion of ortho- to para-hydrogen is of the order of a day or more. Normal warm H_2, having three ortho spin states and one para spin state is three-quarters ortho-hydrogen. On cooling reasonably rapidly to low temperatures the fraction of three-quarters ortho-hydrogen persists; the slow conversion to the equilibrium fraction of practically unity for para-hydrogen goes as a slow first order chemical reaction. Since the energy difference, $\varepsilon/k = 169.9°\text{K}$, of the reactants and products is large, considerable heat is evolved.

One may thus treat the computation of the thermodynamic properties of H_2 gas in two different extremes, depending on the experimental conditions.

In one extreme, corresponding to making slow changes of state in the presence of a good magnetic catalyst, one computes the rotational partition function by summing over all j-values, weighting the odd states with the factor three for the three spin states of ortho-hydrogen. In the other approximation, corresponding to experiments made in a short time in the absence of a catalyst one treats the gas as a mixture of two separate chemically identifiable species, ortho- and para-hydrogen, having mol fraction x_0, of the ortho species. These two species have significantly different energies at zero temperature. In ortho-hydrogen only the odd rotational levels occur, and the nuclear spin entropy per molecule is $k\ln 3$. In para-hydrogen only even j values are allowed and the nuclear spin entropy is zero.

For both species the first excited rotational level has a very high energy. In para-hydrogen the first excited state of $j = 2$ has $\varepsilon/k = 6B_0hc/k = 509.8°\text{K}$; in ortho-hydrogen the energy difference between $j = 3$ and the lowest state of $j = 1$ is $10B_0hc$, $E/k = 849.7°\text{K}$. Even at room temperature the heat capacity due to rotation is not completely classical, and at 20°K it is effectively zero.

For a molecule consisting of two isotopes, HD, HT, or DT, there is no symmetry requirement, since permutation of the two different nucleii not only permutes protons and neutrons but also displaces an extra neutron. All rotational states exist with each of the $(2s_1 + 1)(2s_2 + 1)$ different nuclear spin states.

In the molecule D_2 each deuterium nucleus has spin unity so that there are three states for each nucleus, $\alpha_1, \alpha_2, \alpha_3$, or in the general case $2s + 1$ states, $\alpha_1, \ldots, \alpha_{2s-1}$. The $2s + 1$ combinations $\alpha_\nu(i)\ \alpha_\nu(j)$ are necessarily symmetric, whereas from the remaining $(2s + 1)^2 - (2s + 1)$ products $\alpha_\nu(i)\alpha_\mu(j)$ with $\nu \neq \mu$ one can make $\frac{1}{2}[(2s + 1)^2 - (2s + 1)]$ symmetric and as many unsymmetric spin states, or in all

$$\begin{aligned} &\tfrac{1}{2}(2s + 2)(2s + 1) \text{ symmetric spin states,} \\ &\tfrac{1}{2}(2s(2s + 1) \text{ antisymmetric spin states.} \end{aligned} \tag{3.5.2}$$

If the nuclear mass number is even there are an even number of protons plus neutrons in the nucleus, and s is an integer. In this case the total state function must be symmetric in nuclear exchange, and the lowest rotational level of $j = 0$ will exist with the $\frac{1}{2}(2s + 2)(2s + 1)$ symmetric spin states. Thus D_2 with $s = 1$ has 6 symmetric nuclear spin states in each even rotational state, and 3 antisymmetric nuclear spin states in each antisymmetric rotational state of odd j-value.

In general, for polyatomic molecules of symmetry number σ any σ consecutive rotational states will have differing symmetry properties, each occurring only with its own fraction of the $\Pi(2s_i + 1)$ different nuclear spin states, so that the average number of spin states per single j-level is $\sigma^{-1}\Pi(2s_i + 1)$.

3.6. Polyatomic molecules

For a molecule consisting of n atoms one could use the $3n$ cartesian coordinates, $x_1, y_1, z_1, x_2, \ldots, z_n$ of the atoms to define its configuration. Actually one chooses three cartesian coordinates of the center of mass, $X = (\sum_i m_i)^{-1} \sum_{i=1}^{i=n} m_i x_i$, Y, Z, and $3n - 3$ internal coordinates. If the equilibrium position of minimum potential energy is linear then two angles, the θ and ϕ used for the diatomic molecule axis, are used as rotational coordinates and there remain $3n - 5$ vibrational coordinates. If the equilibrium configuration is non-linear there are *three* angles necessary to specify the rotational configuration, and there remain only $3n - 6$ vibrational coordinates upon which the potential energy depends. As with the diatomic molecule one develops the potential energy as a power series in the displacements from the equilibrium configuration, and since this is the configuration for which the linear terms disappear, the series, after a constant value, starts with only quadratic terms. Higher terms are discarded.

It is then always possible† to choose a transformation of the coordinates so that the potential energy contains only square terms, $\sum_\alpha \frac{1}{2}\kappa_\alpha q_\alpha^2$, and no cross products, $\kappa_{\alpha\beta}q_\alpha q_\beta$, whereas the kinetic energy also remains in the form $\sum_\alpha \frac{1}{2}\mu_\alpha \dot{q}_\alpha^2$ with μ_α a reduced mass. The translational kinetic energy is independent of the internal coordinates. The rotational energy is purely kinetic, and if the moment of inertia of the equilibrium configuration is substituted for the true value it is then independent of the vibrational coordinates. One has now a separable Hamiltonian,

$$H(\boldsymbol{p}^{(3n)}, \boldsymbol{q}^{(3n)}) = H_{\text{trans}} + H_{\text{rot}} + \sum_{\alpha=1}^{3n-(5\text{ or }6)} H_\alpha, \qquad (3.6.1)$$

with

$$H_\alpha = \frac{1}{2\mu_\alpha} p_\alpha^2 + \tfrac{1}{2}\kappa_\alpha q_\alpha, \qquad (3.6.2)$$

of the same quadratic form as the vibrational Hamiltonian for the diatomic molecule, and with the same quantum mechanical stationary states of energy

$$\varepsilon_\alpha(n_\alpha) = n_\alpha h\nu_\alpha,$$
$$\nu_\alpha = (1/2\pi)\sqrt{(\kappa_\alpha/\mu_\alpha)}, \qquad (3.6.3)$$

above the lowest energy level of $\frac{1}{2}h\nu_\alpha$. Finally the internal partition function Q_i is a product,

$$Q_i = Q_{\text{rot}} \prod_{\alpha=1}^{3n-(5\text{ or }6)} Q_\alpha, \qquad (3.6.4)$$

and its logarithm a sum,

$$\ln Q_i = \ln Q_{\text{rot}} + \sum_{\alpha=1}^{3n-(5\text{ or }6)} \ln Q_\alpha, \qquad (3.6.5)$$

with corresponding additive contributions to the chemical potential, μ.

---

† The details of the procedure are carried out for crystals in section 5.2.

The rotational partition function, Q_{rot}, in (4) is now different for the cases that the equilibrium configuration of the molecule is linear or non-linear. If linear, it is the same as that of the diatomic molecule, eqs. (3.4.4), (3.4.5) and (3.4.6) but with the moment of inertia, I_0, given by

$$I_0 = \sum_{i=1}^{i=n} m_i(x_i - X)^2 \tag{3.6.6}$$

$$X = (\sum_{i=n}^{i-n} m_i)^{-1} \sum_{i=1}^{i=1} m_i x_i \tag{3.6.7}$$

which gives, with $M = \Sigma m_i$,

$$I_0 = \frac{1}{M} \sum_{n \geqslant i > j \geqslant 1} \sum m_i m_j (x_i - x_j)^2 \tag{3.6.8}$$

where the positions, x_i, are equilibrium positions of the atom i when the molecule lies along the x axis.

3.7. Polyatomic rotator function

The non-linear polyatomic molecule has a more complicated rotational Hamiltonian. Actually the energies of the allowed quantum levels of rotation in the more general case are not simple and there is no closed analytical solution to their values. However, the moments of inertia of polyatomic molecules are generally larger than those of diatomic molecules, and the gas range is at a higher temperature, so that a purely classical treatment suffices. Fortunately although the classical Hamiltonian looks rather formidable, analytical integration of the expression exp $-(1/kT)H$ can be performed and leads to a quite simple answer.

In any cartesian frame of reference let the positions at equilibrium of atom i be given by x_i, y_i, z_i. The moments of inertia about the cartesian axes passing through the center of mass are, with $M = \sum_i m_i$,

$$I_{xx} = (M)^{-} \sum_{n \geqslant i > j \geqslant 1} \sum m_i m_j [(z_i - z_j)^2 + (y_i - y_j)^2] \tag{3.7.1}$$

with corresponding expressions for I_{yy} and I_{zz}. The products of inertia are I_{yz}, I_{zx} and I_{xy} with

$$I_{yz} = (M)^{-1} \sum \sum_j m_i m_j (x_i - x_j)(y_i - y_j) = I_{zy}. \tag{3.7.2}$$

Four equations

$$\left.\begin{aligned} \alpha(I_{xx}-\eta)-\beta I_{xy}-\gamma I_{xz}&=0,\\ -\alpha I_{yx}+\beta(I_{yy}-\eta)-\gamma I_{yz}&=0,\\ -\alpha I_{zx}-\beta I_{zy}+\gamma(I_{zz}-\eta)&=0, \end{aligned}\right\} \tag{3.7.3}$$

and

$$\alpha^2+\beta^2+\gamma^2=1 \tag{3.7.4}$$

can be solved for η, α, β, γ, with three generally different values. The three values of $\eta = A$, B, and C, are the three principal moments of inertia. The corresponding values of α, β, γ, are the direction cosines of the three principal axes. If the original choice of cartesian coordinates were so made that all three products of inertia were zero, then the three moments, I_{xx}, I_{yy}, I_{zz}, would have been the principal moments. These moments are such that for displacements of the angle of the axis through the center of mass (if they are all different) one is a maximum, another a minimum and the third is at a saddle point.

The three Eulerian angles, θ, ϕ, and ψ, are used to describe the orientation of a rigid body in space.

If x, y, and z are taken as the three Cartesian coordinates fixed in space, and ξ, η, ζ as the three principal (perpendicular) axes of the body, then θ is the angle between the body axis ζ and the space axis z. The line in the xy plane perpendicular to the plane common to ζ and z is called the nodal line. The angle between this nodal line and the x axis is ϕ. Thus θ and ϕ completely determine the direction of the ζ axis in space. The angle between the nodal line and the ξ axis in the body is ψ. This, then, completely defines the orientation of the whole body with fixed center of gravity.

The angle θ may vary between 0 and π; the two angles ϕ and ψ take all values from 0 to 2π.

The Hamiltonian of the rigid body with fixed center of gravity is just the kinetic energy, written as a function of these angles and their conjugated momenta, namely,

$$H=\frac{\sin^2\psi}{2A}\left\{p_\theta-\frac{\cos\psi}{\sin\theta\sin\psi}(p_\phi-\cos\theta\, p_\psi)\right\}^2$$

$$+\frac{\cos^2\psi}{2B}\left\{p_\theta+\frac{\sin^2\psi}{\sin\theta\cos\psi}(p_\phi-\cos\theta p_\psi)\right\}^2+\frac{1}{2C}p_\psi{}^2. \tag{3.7.5}$$

This can be transformed into an expression which will be found more convenient for future operations,

$$\frac{H}{kT} = \frac{1}{2kT}\left(\frac{\sin^2\psi}{A}+\frac{\cos^2\psi}{B}\right)$$
$$\left\{p_\theta + \left(\frac{1}{B}-\frac{1}{A}\right)\frac{\sin\psi\cos\psi}{\sin\theta\left(\dfrac{\sin^2\psi}{A}+\dfrac{\cos^2\psi}{B}\right)}(p_\phi - \cos\theta p_\psi)\right\}^2$$
$$+\frac{1}{2kTAB\sin^2\theta}\frac{1}{\left(\dfrac{\sin^2\psi}{A}+\dfrac{\cos^2\psi}{B}\right)}(p_\phi - \cos\theta p_\psi)^2 + \frac{1}{2kTC}p_\psi^2. \tag{3.7.6}$$

The partition function due to the rotation of this body is

$$Q_{\text{rot}} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{0}^{\pi}\int_{0}^{2\pi}\int_{0}^{2\pi}\frac{1}{h^3}e^{-H(p,q)/kT}\,dp_\theta dp_\phi dp_\psi d\theta d\phi d\psi.$$

The substitution of (5) and (6) appears to lead to a rather formidable integral, but direct integration in the order p_θ, p_ϕ, p_ψ actually offers no difficulties. It is necessary to remember that

$$\int_{-\infty}^{+\infty} e^{-a(x+b)^2}dx = \int_{-\infty}^{+\infty} e^{-ax^2}dx = \left(\frac{\pi}{a}\right)^{1/2}.$$

Integration over p_θ leads to the factor

$$(2\pi kT)^{1/2}\left\{\frac{\sin^2\psi}{A}+\frac{\cos^2\psi}{B}\right\}^{-1/2}.$$

Subsequent integration over p_ϕ yields

$$(2\pi kT\,AB)^{1/2}\sin\theta\left\{\frac{\sin^2\psi}{A}+\frac{\cos^2\psi}{B}\right\}^{1/2},$$

as a factor, which cancels part of that obtained in the first integration. Integration over p_ψ yields the factor

$$(2\pi kTC)^{1/2}.$$

Integration of $\sin\theta\, d\theta$ from 0 to π gives 2, and the other angles give each a factor 2π, so that by defining

$$\xi_A = (h^2/8\pi^2 AkT),\ \text{etc.}, \tag{3.7.7}$$

in conformity with (3.4.4) for the diatomic molecule, we arrive at

$$Q_{ro} = (\pi/\xi_A\xi_B\xi_C)^{1/2} \tag{3.7.8}$$

for a non-linear polyatomic molecule.

Since Q_{rot} is proportional to $T^{3/2}$ one obtains the classical limit of $(3/2)kT$ for the contribution of the three degrees of freedom to the energy per molecule due to the kinetic energy alone, section 2.4.

3.8. Polyatomic normal coordinates

The coordinates for which the quadratic approximation to the potential energy is a sum of perfect squares, $u = \Sigma\frac{1}{2}\kappa_\alpha q_\alpha^2$, are called normal coordinates. Their determination for complicated polyatomic molecules constitutes a field of scientific endeavor by itself. Only a few comments pertinent to their use in partition functions need concern us here.

As is the case with diatomic molecules the purely theoretical computation of the force constants from *a priori* quantum mechanical equations is usually both very difficult and inaccurate. Recourse is had to spectroscopically measured frequencies. If all $3n - 5$ or $3n - 6$ different vibration frequencies are observed the computation of the partition function can be made without any further reference to the nature of the coordinate transformation. This is seldom the case.

If the molecule under consideration has any non-trivial symmetry the use of group theory facilitates the analysis. (If the molecule is simple, with simple symmetry, the group theory is still valid, but the answer is often obvious without it.) One then finds that certain frequencies are degenerate, that is, there exist sets of two or more normal coordinates along which the vibrations have identical frequencies. In such a case, of course, only fewer than the theoretical $3n - (5 \text{ or } 6)$ frequencies can be experimentally observed, and one must decide which ones to count the appropriate number of times in summing $\ln Q_\alpha$.

The group theoretical analysis gives, with the symmetry class of the coordinate, also selection rules: whether the frequency will be observed in the infrared or in the Raman, or in neither or both, and whether it should occur in combination with electronic transitions of given symmetry. If sufficient spectroscopic data is available including the overtones this may suffice to identify the different observed values with coordinates of certain symmetries having known degeneracies.

Very often at least one, or sometimes more of the frequencies are unknown, and must be estimated from the values of the others. No rigorous method for doing this exists, but various approximations appropriate to particular cases can be made, and are sometimes convincingly plausible.

3.9. Polyatomic molecule symmetries

In assigning symmetry numbers to polyatomic molecules cases occasionally arise in which the choice appears to be doubtful or arbitrary. Since the symmetry number enters the entropy per molecule as $-k \ln \sigma$ one might well be dismayed that an infinitesimal difference in the position of the equilibrium configuration could cause a discontinuous change in the entropy. A similar situation enters in the case of a linear polyatomic molecule with regard to the assignment of the numbers of rotational and vibrational degrees of freedom, since a molecule infinitesimally off linear in its equilibrium position now acquires one extra rotational and one less vibrational degree of freedom. Careful examination of each case reassures one that nothing is amiss. A few examples will be discussed, especially since they point up serious limitations on the use of the analytical approximations of the last few sections.

Consider the case of ammonia, NH_3, and ammonia-like molecules MX_3, in which the three atoms X are at the vertices of an equilateral triangle and the central atom M is on a line normal to the plane of the three X-atoms, passing through the center of the triangle. If the equilibrium position, z_0, of M with respect to displacement z along this line is at $z_0 = 0$ in the plane of the three X's then the symmetry number is 6, but if it is off the plane then the symmetry number is 3. The question arises as to the treatment when z_0 is small. The answer lies in the obvious symmetry that if $z_0 \neq 0$ there must be two symmetrical minima, u_{min} of the potential $u(z)$, namely at z_0 and at $-z_0$, with presumably a maximum, u_{max} at $z = 0$. Now if z_0 is large, with $u_{\text{max}} - u_{\text{min}} >> kT$, as well as $u_{\text{max}} - u_{\text{min}} >> h\nu$, then a treatment using *one* minimum alone with $\sigma = 3$ and a harmonic approximation, $u(z) = u_{\text{min}} + \frac{1}{2}\kappa(z - z_0)^2$ will be adequate. However, one could alternatively consider z to go from $-\infty$ to $+\infty$ with the two minima and solve the double minimum quantum mechanical problem for the energies ε_n of the states. Then sum $\exp - (\varepsilon_n/kT)$ for these states, and use $\sigma = 6$. If the minima are well separated the solution for the double minimum energies will be that of paired states very close in energy to each of the low lying single minimum states. The sum of $\exp - (\varepsilon_n/kT)$ will be twice as great, and division by six will lead to the single minimum result divided by three.

If, however, the minimum positions z_0 and $-z_0$ lie close to zero, so that $u_{\max} - u_{\min}$ will be small, the double minima spectra will no longer resemble that of the single minimum harmonic approximation. A numerical solution of the quantum levels must be made and the sum of $\exp - (\varepsilon_n/kT)$ made numerically, using $\sigma = 6$. Finally in the limit that the two minima fuse at $z_0 = -z_0 = 0$ one may again expect the harmonic approximation with $\sigma = 6$ to be valid.

A somewhat analogous situation exists for ethane, $H_3C—CH_3$. One internal coordinate is an angle χ running from 0 to 2π measuring equal and opposite rotational displacements of the two CH_3 groups around their common axis. Two extreme situations lend themselves to simple analytical approximation. The first is that the triply periodic potential, $u(\chi)$ has minima separated by high maxima. Use of a harmonic potential development about any one minimum with $\sigma = 6$ may be expected to give a good approximation. The other limit is that $u(\chi)$ is constant. The states would then be states of free rotation having kinetic energy only given by $\varepsilon_n = (h^2/16\pi^2C)n^2$ with C the moment of inertia of a single CH_3 group and n an integer. The sum of $\exp - (\varepsilon_n/kT)$ must be divided by $\sigma = 18$ in this case, since 18 different configurations of numbered identical molecules are covered in the integration. The more realistic intermediate case of the triply periodic potential with relatively low maxima requires numerical solution of the triple minima quantum mechanical problem and then numerical solution of the partition function sum with the energies obtained.

The case of the almost linear molecule, say one of type XMX represents a similar discontinuity in language, although the symmetry number remains two, whether the molecule is linear or not. The linear molecule has only two degrees of rotational freedom and $3n - 5$ or four vibrational degrees of freedom. Two of these correspond to two vibrations along the axis of the molecule which will have different frequencies. One of these has the two X's moving together against the counter motion of the central M atom. In the other the M is stationary and the two X atoms vibrate in and out against it. The two remaining normal vibrational coordinates are bending along the two perpendicular directions in the plane normal to the molecular axis. These, from symmetry, have necessarily equal frequencies, and are degenerate, so that linear combinations of their excited states are also stationary states of allowed energy.

Now among the linear combinations are those which correspond to classical rotation of the bent molecule with fixed angle of deformation.

The rotation is about a principal axis of the molecule lying in the plane of the three atoms, and parallel to the line joining the two X atoms, with the centrifugal force due to rotation just balancing the restoring force. For higher excitations there are linear combinations corresponding to "end on" pictures of motion in an ellipse, combining rotation and vibration. Thus vibrational states of these two degrees of freedom do include states having angular momenta about the axis of the molecule. There will always be paired states of equal energy having equal and opposite angular momenta.

Now the molecule having a slightly bent equilibrium configuration would be represented as having only one vibrational degree of freedom due to bending in the plane of the three atoms, and one new rotational degree of freedom about an axis in this plane, an axis which will be the principal axis of the molecule having the lowest moment of inertia. Since the moment of inertia is low, rotational states about this axis will have a high energy, comparable with that of a bond-bending vibrational state. Since the rotational energies of even the lower excited states would be so high the centrifugal forces would be large. Also the moment of inertia about this axis would be critically dependent on the vibrational bending coordinate, so that the separability of the Hamiltonian into a purely rotational and purely vibrational term would be illusory, and the quantum mechanical energy expression would have important terms involving both rotational and vibrational quantum numbers. No simple independent additive contributions from rotation and vibration to $\ln Q_i$ can be expected. A complete new quantum mechanical analysis of motion in an r–σ plane with a potential minimum in a ring of small r-values is required, and no discontinuity in behavior occurs when the ring shrinks to a point at $r = 0$.

The simple analysis in terms of normal coordinates for which the Hamiltonian is separable into pure rotational and independent harmonic oscillations is always an approximation whose validity must be examined for individual cases. In those cases where an infinitesimal variation of some parameter from a unique symmetry position would appear to introduce a discontinuous change in values the approximation is utterly invalid. When the approximation is still reasonably valid, or when it is quite good but very high accuracy of numerical results is desired there are perturbation procedures† by which the terms in the quantum mechanical energy expressions proportional to n^2, to $j^2(j + 1)^2$

---

† See, for instance, Chapter 7 in Mayer and Mayer, *Statistical Mechanics*, John Wiley and Sons, 1940.

and to $nj(j + 1)$ can be incorporated into the partition function sums when the coefficients are known.

However, modern computer techniques make increasingly attractive purely numerical summations when energy level values are known, as they also make possible increasingly reliable quantum mechanical energy level calculations for those cases in which no analytical expressions occur.

3.10. Perfect gas thermodynamic functions

We may summarize the results of this chapter by presenting a list of the equations for the common thermodynamic functions of a perfect gas. We use $G = N\mu$, $(\partial G/\partial P)_{N,T} = V$ and $(\partial G/\partial T)_{N,P} = -S$. With these $A = G - PV$, $H = G + TS$ and $E = H - PV$. Finally, we list $C_P = (\partial H/\partial T)_P$ and $C_V = (\partial E/\partial T)_V$. All these functions can be written as sums of contributions from the translational and internal degrees of freedom, respectively, and in many cases with satisfactory precision the internal contribution can be given as a sum of terms due to rotation and single vibrational contributions. The equations are given per mol using $R = kN_0$ the gas constant per mole with N_0 Avogadro's number.

3.10a. *The translational contributions, X_i, per mol*

With m the molecular mass, and M the molecular weight,

$$\begin{aligned}
\lambda &= (h^2/2\pi mkT)^{1/2} \\
&= (1.745\,72 \times 10^{-7})M^{-1/2}T^{1/2} \\
&= \sqrt{(300/MT)}\text{ångstrom}, \\
G_k &= RT \ln \lambda^3(N_0/V) \\
&= RT \ln \lambda^3(P/kT) \\
&= RT\left(-\frac{3}{2}\ln M - \frac{3}{2}\ln T - \ln V + K_{GV}\right) \\
&= RT\left(-\frac{3}{2}\ln M - \frac{5}{2}\ln T + \ln P + K_{GP}\right),
\end{aligned}$$

$$\begin{aligned} A_k &= RT \ln \lambda^3 (N_0/V) e^{-1} \\ &= RT \ln \lambda^3 (P/kT) e^{-1} \\ &= RT\left(-\frac{3}{2}\ln M - \frac{3}{2}\ln T - \ln V + K_{AV}\right) \\ &= RT\left(-\frac{3}{2}\ln M - \frac{5}{2}\ln T + \ln P + K_{AP}\right), \end{aligned}$$

$$\begin{aligned} S_k &= R \ln \lambda^{-3} (V/N_0) e^{5/2} \\ &= R \ln \lambda^{-3} (kT/P) e^{5/2} \\ &= R\left[\frac{3}{2}\ln M + \frac{3}{2}\ln T + \ln V + K_{SV}\right] \\ &= R\left[\frac{3}{2}\ln M + \frac{5}{2}\ln T - \ln P + K_{SP}\right], \end{aligned}$$

$$H_k = \frac{5}{2}RT, \quad E_k = \frac{3}{2}RT,$$

$$C_{Pk} = \frac{5}{2}R, \quad C_{Vk} = \frac{3}{2}R.$$

The present values of the constants occurring here are,

$N_0 = (6.022\,52 \pm 0.000\,28) \times 10^{23}\ \mathrm{mol}^{-1}$

$h = (6.625\,6 \pm 0.000\,5) \times 10^{-27}$ erg sec,

$k = (1.380\,54 \pm 0.000\,18) \times 10^{-16}$ erg deg^{-1},

$R = (8.314\,3 \pm 0.001\,2) \times 10^{7}$ erg $\mathrm{mol}^{-1}\ \mathrm{deg}^{-1}$,

and conversion factors are,

1 atm = 1.013 249 bars = 760 mm,

1 cal = 4.184 0 joule = $4.184\,0 \times 10^{7}$ erg.

With these and $\ln x = 2.302\,585 \log_{10} x$ we have the values in Tables 3.1 and 3.2.

TABLE 3.1

Quantity	joules mol^{-1} deg^{-1}	cal mol^{-1} deg^{-1}
R	8.314 3	1.987 2
$\frac{1}{2}R \ln x$	9.572 2 $\log_{10} x$	2.287 8 $\log_{10} x$
$R \ln x$	19.144 4 $\log_{10} x$	4.575 6 $\log_{10} x$
$\frac{3}{2}R \ln x$	28.716 6 $\log_{10} x$	6.863 4 $\log_{10} x$
$2R \ln x$	38.288 8 $\log_{10} x$	9.151 2 $\log_{10} x$
$\frac{5}{2}R \ln x$	47.861 0 $\log_{10} x$	11.439 0 $\log_{10} x$

TABLE 3.2

Constant	Units	Value of constant	$R \times$ constant	
			joule mol^{-1} deg^{-1}	cal^{-1} mol^{-1} deg^{-1}
K_{GV}	V cc	8.072 7	67.118 8	16.041 8
K_{GV}	V liters	1.164 9	9.685 3	2.314 8
K_{GP}	P atm	3.665 0	30.471 9	7.283 0
K_{GP}	P mm	−2.968 3	−24.679 3	−5.898 5
K_{GP}	P bar	3.651 9	30.363 0	7.256 9
K_{AV}	V cc	7.072 7	58.804 5	14.054 6
K_{AV}	V liters	0.164 9	1.371 0	0.327 7
K_{AP}	P atm	2.665 0	22.157 6	5.295 8
K_{AP}	P mm	−3.968 3	32.993 6	−7.885 7
K_{AP}	P bar	2.651 9	22.048 7	5.269 8
K_{SV}	V cc	−5.572 7	46.333 1	−11.073 9
K_{SV}	V liters	1.335 1	11.100 4	2.653 1
K_{SP}	P atm	−1.165 0	9.686 2	−2.315 0
K_{SP}	P mm	5.468 3	45.465 1	10.866 4
K_{SP}	P bar	−1.151 9	9.577 2	−2.289 0

3.10*b*. *The internal contributions, X_i per mol*

With Q_i the internal partition function,

$$Q_i = \sum_{\substack{\text{all internal} \\ \text{states } i}} e^{-\varepsilon_i/kT}, \quad W_i = Q_i^{-1}\, e^{-\varepsilon_i/kT},$$

we have

$$G_i = A_i = -RT \ln Q_i,$$

$$S_i = R \frac{d}{dT} (T \ln Q_i)$$

$$= R \sum_i \left[\ln Q_i + \frac{\varepsilon_i}{kT} \right] [Q_i^{-1} e^{-\varepsilon_i/kT}]$$

$$= -R \sum_i W_i \ln W_i,$$

$$H_i = E_i = RT^2 \frac{d}{dT} \ln Q_i$$

$$= N_0 Q_i^{-1} \sum \varepsilon_i e^{-\varepsilon_i/kT} = N_0 < \varepsilon_i >,$$

$$C_{Pi} = C_{Vi} = R \frac{d}{dT} T^2 \frac{d}{dT} \ln Q_i$$

$$= R Q_i^{-2} \sum_i \sum_j \left(\frac{\varepsilon_i}{kT} - \frac{\varepsilon_i}{kT} \right)^2 e^{-(\varepsilon_i + \varepsilon_j)/kT}$$

$$= (N_0/kT^2)(< \varepsilon_i^2 > - < \varepsilon_i >^2).$$

3.10c. *High temperature nuclear spin contributions*

Except for H_2 and some hydrogen derivatives temperatures for which the materials are gaseous may usually be regarded as sufficiently high. The contribution due to nuclear spin magnitude S_n, *per gram atom of each element* is

$$G_{\text{spin}} = A_{\text{spin}} = -RT \ln (2S_n + 1),$$
$$S_{\text{spin}} = R \ln (2S_n + 1),$$
$$H_{\text{spin}} = E_{\text{spin}} = C_{P_{\text{spin}}} = C_{V_{\text{spin}}} = 0.$$

3.10d. *Heavy element isotope mixing contribution*

If x_j is the fraction of isotope j, and if all mechanical effects such as change of vibrational frequencies due to the difference of masses of the different isotopes can be neglected, then due to the non-identity of the atoms there is *per gram atom of each element* a contribution,

$$G_{\text{mix}} = A_{\text{mix}} = RT \sum_j x_j \ln x_j,$$
$$S_{\text{mix}} = -R \sum_j x_j \ln x_j,$$
$$H_{\text{mix}} = E_{\text{mix}} = C_{P_{\text{mix}}} = C_{V_{\text{mix}}} = 0.$$

3.10e. *Monatomic gas internal contributions*

In addition to the nuclear spin and isotope mixing the sum $\Sigma e^{-\varepsilon_i/kT}$ must be made over the electronic states. Most frequently below several thousand degrees only the lowest level of energy $\varepsilon_0 = 0$ need be counted, and if this is g-fold degenerate, $Q_{\text{electronic}} = g$,

$$G_{\text{elec}} = A_{\text{elec}} = -RT \ln g,\ S_{\text{elec}} = R \ln g,$$

$$H_{\text{elec}} = E_{\text{elec}} = C_{P_{\text{elec}}} = C_{V_{\text{elec}}} = 0.$$

3.10f. *Rotational contributions*

For molecules that are reasonably rigid in structure the Hamiltonian can be written as a sum of a rotational and an independent vibrational term. Insofar as this is justified there is an additive rotational contribution to the thermodynamic functions.

If the molecule is linear, either diatomic or linear polyatomic, with moment of inertia I_0, the quantity $B_0 = h/8\pi^2 I_0 c$ is often listed in cm^{-1} in spectroscopic tables. With

$$\xi = (h^2/8\pi^2 I_0 kT) = B_0 hc/kT,$$

the contributions due to the rotation of the rigid linear molecule are, when $\xi < 2/3$,

$$G_{\text{rot}} = A_{\text{rot}} = RT\left[\ln \xi - \frac{1}{3}\xi - \frac{1}{90}\xi^2 - \frac{8}{2835}\xi^3 - \cdots\right]$$

$$S_{\text{rot}} = R\left[1 - \ln \xi - \frac{1}{90}\xi^2 - \frac{16}{2835}\xi^3 - \cdots\right]$$

$$H_{\text{rot}} = E_{\text{rot}} = RT\left[1 - \frac{1}{3}\xi - \frac{1}{45}\xi^2 - \frac{8}{945}\xi^3 - \cdots\right]$$

$$C_{P_{\text{rot}}} = C_{V_{\text{rot}}} = R\left[1 + \frac{1}{45}\xi^2 + \frac{16}{945}\xi^3 + \cdots\right]$$

If the molecule is symmetrical upon rotation by π, the symmetry contributions with $\sigma = 2$ must be added. These are

$$G_{\text{sym}} = A_{\text{sym}} = +RT \ln \sigma,$$

$$S_{\text{sym}} = -R \ln \sigma,$$

$$H_{\text{sym}} = E_{\text{sym}} = C_{P_{\text{sym}}} = C_{V_{\text{sym}}} = 0.$$

A polyatomic molecule having a non-linear equilibrium configuration has three principal moments of inertia, A, B, C. With

$$\xi_A = (h^2/8\pi^2 AkT), \text{ etc.},$$

the rotational contribution is then

$$G_{\text{rot}} = A_{\text{rot}} = \frac{1}{2} RT \ln \pi^{-1} \xi_A \xi_B \xi_C,$$

$$S_{\text{rot}} = \frac{1}{2} R[\ln e^3 \pi(\xi_A \xi_B \xi_C)^{-1}],$$

$$H_{\text{rot}} = E_{\text{rot}} = \frac{3}{2} RT,$$

$$C_{P_{\text{rot}}} = C_{V_{\text{rot}}} = \frac{3}{2} R,$$

plus the symmetry number contributions listed above.

3.10*g*. *Vibrational contributions*

For one mol of molecules each vibrational degree of freedom of frequency ν contributes to the thermodynamic functions an additive term as follows:

$$G_{\text{vib}} = A_{\text{vib}} = RT \ln [1 - e^{-h\nu/kT}],$$

$$S_{\text{vib}} = R[(h\nu/kT)(e^{h\nu/kT} - 1)^{-1} - \ln (1 - e^{-h\nu/kT})],$$

$$H_{\text{vib}} = E_{\text{vib}} = RT(h\nu/kT)(e^{h\nu/kT} - 1)^{-1},$$

$$C_{P_{\text{vib}}} = C_{V_{\text{vib}}} = R(h\nu/kT)e^{h\nu/kT}(e^{h\nu/kT} - 1)^{-2}.$$

CHAPTER 4

IMPERFECT GASES

4.1. Introduction

The treatment of imperfect gases might well be included with the later chapter on dense fluids. The particular method which is most useful in treating imperfect gases has considerable usefulness, as a general mathematical tool, in the treatment of dense fluids, including liquids. Its physical significance is so much clearer in the treatment of gases that it appears to be useful to present it, in its simplest form, first in the interpretation of this case.

The method is intuitively simple. We consider the case of a one-component system, and for simplicity, monatomic molecules with no internal electronic excited levels of energy comparable with kT. We assume the classical Boltzmann limit to be valid, and use (2.4.5) for Q_c,

$$Q_c(N,V,T) = \lambda^{-3N} Q_\tau / N!, \tag{4.1.1}$$

$$Q_\tau(N,V,T) = \int\int_V \dots \int e^{-U^N/kT} dx_i \dots dz_N. \tag{4.1.2}$$

Define an activity z as

$$z = \lambda^{-3} e^{\mu/kT}. \tag{4.1.3}$$

From eq. (3.1.1) with $Q_i = 1$, in the perfect gas limit $z = P/kT$, which in turn equals

$$\rho = \bar{N}/V, \tag{4.1.4}$$

namely the number density, and

$$\lim_{z=0} \left[\frac{z}{\rho}\right] = 1. \quad \text{(Perfect Gas)} \tag{4.1.5}$$

The grand canonical partition is, then, from (1.4.3) and (1.4.5), with (1) and (3)

$$e^{PV/kT} = Q_{G\cdot C} = \sum_{n>0} \frac{z^N}{N!} Q_\tau(N,V,T). \tag{4.1.6}$$

The perfect gas limit follows immediately if $U_N = 0$, in which case the integrand of (2) is everywhere unity, and $Q_\tau = V^N$, $\Sigma(zV)^N/N! = e^{Vz}$, $P/kT = z$.

In the general case the integrand $e^{-U/kT}$ of (2) is not everywhere unity. We make use of the fact that it is unity when all pair distances, r_{ij}, are large, and write it as one plus correction terms which are non-zero only when clusters of molecules are close together. These correction terms can be written as products of functions of the coordinates of small subsets of molecules, so that the integral over them is a product. The number of terms of a given product is such that P/kT itself becomes a power series in z; the coefficient of z^n is the cluster integral over a cluster of n molecules.

4.2. Simple cluster functions

The structure of the cluster functions is most apparent if the potential energy, U_N, is a sum of pair potentials,

$$U_N = \sum_{N \geqslant i > j \geqslant 1} \sum u(r_{ij}), \tag{4.2.1}$$

so that the exponential is a product,

$$e^{-U_N/kT} = \prod_{N \geqslant i > j \geqslant 1} \prod e^{-u(r_{ij})/kT}. \tag{4.2.2}$$

Write

$$f_{ij} = f(r_{ij}) = [\exp - u(r_{ij})/kT] - 1, \tag{4.2.3}$$

and

$$\begin{aligned} e^{-U_N/kT} = \prod_{n \geqslant i > j \geqslant 1} \prod (1 + f_{ij}) &= 1 + \Sigma\Sigma f_{ij} \\ &+ \Sigma\Sigma\Sigma\Sigma f_{ij} f_{kl} + \ldots \Sigma \ldots \Sigma f_{ij} f_{kl} \ldots f_{mn} + \ldots . \end{aligned} \tag{4.2.4}$$

The general product that occurs in this expansion is most easily visualized by reference to a picture (Fig. 4.1). Draw numbered circles for every one of the N molecules. For every function, f_{ij}, of a given product draw a line between circles i and j. To every product there corresponds one figure, and to every possible figure in which all pairs i, j are either unconnected or have one line between them, there corresponds one product in the sum (4).

The general product may be characterized by the number set, $m_1, m_2, \ldots, m_n$, of clusters of n molecules each. In the corresponding figure there will be m_1 single circles unconnected to any other; m_2 pairs of circles, i, j, connected to each other by a line but to no other circle; m_3 triples, i, j, k, connected in any of four ways, by $f_{ki}f_{ij}$, by $f_{ij}f_{jk}$, by

$f_{jk}f_{ki}$, or by $f_{ij}f_{jk}f_{ki}$, and unconnected to any other; . . .; m_n clusters of n directly or indirectly connected to each other by lines in the figure and by f_{ij}'s in the product, but not connected to others.

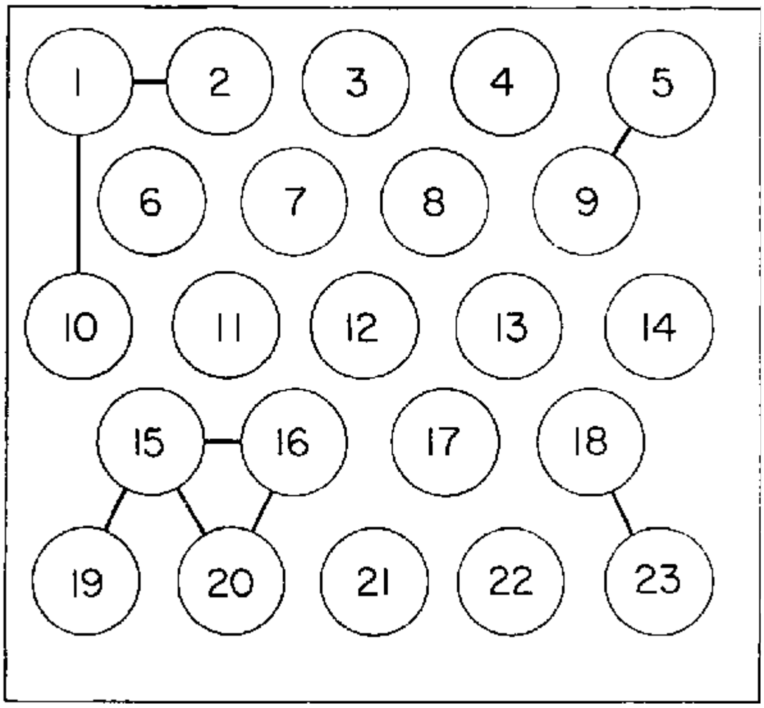

FIG. 4.1. Graphical representation for 23 molecules of the product $f_{1,2}f_{1,10}f_{15,16}f_{15,19}f_{15,20}f_{16,20}f_{18,23}$ in which there are 12 clusters of one molecule each, two clusters of two each, one of three, and one of four.

The cluster function g_n is defined as the sum of all connected products of f_{ij}'s between n molecules.

$$g_n(\mathbf{r}_i, \ldots, \mathbf{r}_n) = \sum_{\alpha_{ij}=0,1} \begin{bmatrix}\text{connected} \\ \text{products}\end{bmatrix} \prod_{n>i>j>1} \prod [f(r_{ij})]^{\alpha_{ij}}, \tag{4.2.5}$$

$$g_1(\mathbf{r}_1) \equiv 1, \tag{4.2.5$'$}$$

and the cluster integrals by

$$b_n(V) = \frac{1}{Vn!} \iint_V \cdots \int g_n d\mathbf{r}_1 \ldots d\mathbf{r}_n, \tag{4.2.6}$$

$$b_1 \equiv 1. \tag{4.2.6$'$}$$

If the pair potential, $u(r)$, goes to zero more rapidly than r^{-3} for large r, which it does for neutral molecules, the cluster function g_n goes sufficiently rapidly to zero if any of the pair distances approach infinity so that $b_n(V)$ defined by (6) becomes independent of V for large V and finite n We define the limit

$$b_n = \lim_{V\to\infty} [b_n(V)], \tag{4.2.7}$$

which is then finite for finite values of n.

Now return to the sum of products in (4). Sum all possible products in which the same m_1 numbered molecules are in single clusters; the same m_2 numbered pairs are in clusters of 2, . . .; the same m_n clusters of n numbered molecules each are clustered together; . . .. The sum is the product of the cluster functions of these numbered molecules. The integral over V of the sum is the product of their integrals, or

$$\prod_{n>1} (Vn!b_n)^{m_n}, \quad (\sum_{n>1} nm_n = N). \tag{A}$$

For given value of the numbers $m_1, m_2, \ldots, m_n$, there are

$$N!/[\prod_{n\geqslant 1} (n!)^{m_n} m_n!] \tag{B}$$

ways of assigning the N numbered molecules to m_1 clusters of 1 each, . . ., m_n clusters of n each, . . . (see section 2.5). The integral of $Q_\tau/N!$ is then $(A) \times (B)/N!$ summed over all m_n consistent with $\Sigma nm_n = N$,

$$\frac{Q_\tau(N,V,T)}{N!} = \sum_{m_1\geqslant 0} \cdots \sum_{m_n\geqslant 0 (\Sigma nm_n = N)} \prod_{n\geqslant 1} \frac{(Vb_n)^{m_n}}{m_n!}. \tag{4.2.8}$$

Since $\Sigma nm_n = N$ we have $z^N = \prod_{n\geqslant 1} (z^n)^{m_n}$, and,

$$\frac{z^n Q_\tau(N,V,T)}{N!} = \sum_{\{m_n\}, (\Sigma nm_n = N)} \prod_{n\geqslant 1} \frac{(Vb_n z^n)^{m_n}}{m_n!}. \tag{4.2.9}$$

The sum over N now removes the restriction $\Sigma nm_n = N$ on the values that m_n may take. Each m_n may take all values from zero to infinity independently of the values of the others. The sum of products is the product of the sums,

$$\prod_{n\geqslant 1} \sum_{m_n\geqslant 0}^{\infty} \frac{(Vb_n z^n)^{m_n}}{m_n!} = \prod_{n\geqslant 1} \exp Vb_n z^n, \tag{4.2.10}$$

and this is the exponential of the sum. We have from (4.1.2)

$$Q_{G\cdot C} = \exp V \sum_{n\geqslant 1} b_n z^n, \tag{4.2.11}$$

and since $Q_{G\cdot C} = \exp PV/kT$ from (1.4.3),

$$P/kT = \sum_{n\geqslant 1} b_n z^n = z + b_2 z^2 + b_3 z^3 + \cdots, \tag{4.2.12}$$

since $b_1 = 1$, eq. (6′).

4.3. The virial expansion

The next step is to find P/kT as a power series in the number density ρ. The step is completely straightforward, although numerically tedious, if only the coefficients up to some small power of ρ are sought. The result, however, is surprisingly simple, due apparently to a numerical coincidence. The proof that the numerical coincidence is general, and that the series obtained is correct to all powers is more complicated.

The cluster integrals b_n are naturally evaluated as a sum of products of what we call irreducible integrals. The integral b_2 is, from (6),

$$b_2 = (1/2)\int_0^\infty 4\pi r^2 f(r)\, dr. \tag{4.3.1}$$

We define the first irreducible integral, β_1, as

$$\beta_1 = 2b_2 = \int_0^\infty 4\pi r^2 f(r)\, dr. \tag{4.3.2}$$

The cluster function, g_3, for 3 is

$$g_3(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k) = f_{ki}f_{ij} + f_{jk}f_{ki} + f_{ij}f_{jk} + f_{ij}f_{jk}f_{ki}. \tag{4.3.3}$$

Consider the integral of $f_{ki}f_{ij}$. Integrate first over $d\boldsymbol{r}_k$ keeping $\boldsymbol{r}_i\boldsymbol{r}_j$ fixed. One obtains $f_{ij}\beta_1$, and for the total integral $V\beta_1^2$. The terms $f_{ij}f_{jk}$ and $f_{jk}f_{ki}$ give the same integral. Define

$$\beta_2 = \frac{1}{2V}\iiint_V f_{ij}f_{jk}f_{ki}\,d\boldsymbol{r}_i\,d\boldsymbol{r}_j\,d\boldsymbol{r}_k, \tag{4.3.4}$$

one has

$$b_3 = \frac{1}{3\,!}\,[3\beta_1^2 + 2\beta_2]. \tag{4.3.5}$$

Similarly g_4 contains 12 terms corresponding to "chains," $f_{ij}f_{jk}f_{kl}$ and 4 terms of the type $f_{ij}f_{ik}f_{il}$ each of which integrates to β_1^3, as well as 12 terms like $f_{ij}f_{jk}f_{kl}f_{lj}$ which give $2\beta_1\beta_2$ upon integration. Defining β_3 as $1/V3\,!$ times the integral of the remaining terms one has

$$b_4 = \frac{1}{4\,!}\,[16\beta_1^3 + 24\beta_1\beta_2 + 6\beta_3]. \tag{4.3.6}$$

It will turn out that we can write P/kT as

$$P/kT = 1 - \sum_{\nu\geqslant 1}\frac{\nu}{\nu+1}\,\beta_\nu\rho^2. \tag{4.3.7}$$

It is straightforward and easy to check this from (2), (5), and (6) up to $\nu = 3$. From (1.3.12) that $(\partial P/\partial \mu)_T = \bar{N}/V = \rho$, and (4.1.3) defining the activity z, $\mu = kT \ln z\lambda^3$, it follows that

$$[\partial(P/kT)/\partial \ln z]_T = \rho. \tag{4.3.8}$$

Using (4.2.10) with this

$$\rho = \sum_{n \geqslant 1} n b_n z^n = z + 2b_2 z^2 + 3b_3 z^3 + 4b_4 z^4 \ldots. \tag{4.3.9}$$

Write

$$z = \rho + a_1 \rho^2 + a_2 \rho^3 + a_3 \rho^4 + \ldots, \tag{4.3.10}$$

and use this in (9), equating the coefficients of equal powers of ρ to find,

$$\begin{aligned} a_1 &= -2b_2 = -\beta_1, \\ a_2 &= -4b_2 a_1 - 3b_3 = -[\beta_2 - (1/2)\beta_1^2], \\ a_3 &= -[\beta_3 - \beta_1\beta_2 + (1/6)\beta_1^3], \end{aligned} \tag{4.3.11}$$

which, in (10), gives z up to ρ^4 the same as in the expansion of

$$z = \rho \exp - \sum_{\nu \geqslant 1} \beta_\nu \rho^\nu. \tag{4.3.12}$$

We shall later prove (12) to be general. Use the values of a given by (11) in (10) for z, and with these evaluate P/kT from (4.2.10). The answer, up to ρ^4 is that of eq. (7).

We shall now turn to the more difficult problem of proving (12) and (7) to be general.

4.4. Proof of the virial development

The proof consists of the following steps. We define the general irreducible function, γ_ν, of the coordinates of $\nu + 1$ molecules, and the irreducible integral, β_ν, of γ_ν. We then evaluate the total number of terms in the cluster function g_n, which lead to a given product of β_ν's upon integration. This leads to an expression for b_n as a sum of products of β_ν's. From this we derive eq. (4.3.7) and (4.3.12).

Turn to any graph of n numbered vertices, with one line or no line drawn between each pair, i, j, of vertices, in which the whole is a connected graph; a path of lines between successive vertex pairs connecting every pair. Such a graph corresponds to a product in the cluster function g_n, and g_n is the sum of all the products corresponding to all possible graphs drawn according to such a prescription. In the interest of brevity we will identify the product and the graph; speak of a sum of products as a sum of graphs; and the integration over the coordinates of a molecule in the product as integration over the corresponding numbered vertex of the graph.

If, between any pair, k, j of vertices, all paths of consecutive lines necessarily pass through some single vertex α, we call α a nodal vertex, or simply a node of the graph. Each node, α, then, divides the graph into $\lambda_\alpha + 1$ parts, $\lambda_\alpha \geqslant 1$, connected together only by containing each the same nodal vertex α. We call λ_α the "order" of the node α, the non-nodal vertices have order zero (Fig. 4.2).

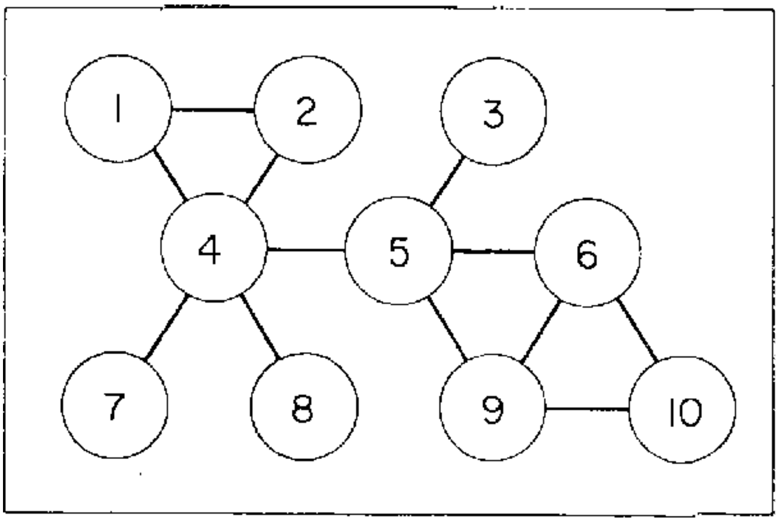

FIG. 4.2. Graph of a term in g_{10} with the nodal vertex four of order 3 and nodal vertex five of order 2. The integral is proportional to $\beta_1{}^3$-β_2 times one term of β_3.

Hold the vertices of all but one part, including the node α, fixed, and integrate over the vertices of the one part. One obtains a factor which is V^{-1} times the integral of that part graph including α over the total volume, V. In this way the integral of the graph can be obtained as V times the product over all the parts between nodes of V^{-1} times their integral. Define the irreducible function γ_ν of the coordinates of $\nu + 1$ molecules as the sum of (products corresponding to) all graphs of $\nu + 1$ molecules containing no nodes:

$$\gamma_\nu(\mathbf{r}_1, \ldots \mathbf{r}_{\nu+1}) = \sum_{\alpha_{ij}=0,1} \begin{bmatrix}\text{sum of connected}\\ \text{products without}\\ \text{nodes}\end{bmatrix} \prod_{\nu+1 \geqslant i > j \geqslant 1} \prod f_{ij}{}^{\alpha_{ij}}, \qquad (4.4.1)$$

which is symmetrical in permutation of the $\nu + 1$ coordinates. The sum of all graphs with the *same* non-nodal vertices in the parts separated by the *same* nodal vertices is then a product of the irreducible functions, γ_ν. Each irreducible function contains the coordinates of at least one nodal vertex, and each nodal vertex α, of order λ_α, occurs in $\lambda_\alpha + 1$ different irreducible functions.

If the sum has, in the product, m_ν irreducible functions γ_ν, then with n the total number of vertices,

$$\sum_{\nu \geqslant 1} (\nu + 1)m_\nu = n + \sum_\alpha \lambda_\alpha. \tag{4.4.2}$$

If

$$\sum_{\nu \geqslant 1} m_\nu = m \tag{4.4.3}$$

is the total number of irreducible functions, then there are $m - 1$ "connections" and since each node, α, of order λ_α makes λ_α connections,

$$\sum_\alpha \lambda_\alpha = m - 1. \tag{4.4.4}$$

With this and (2), we have that

$$\sum_{\nu \geqslant 1} \nu m_\nu = n - 1. \tag{4.4.5}$$

The irreducible integral, β_ν, is defined as

$$\beta_\nu = \frac{1}{V\nu!} \int\int_V \cdots \int \gamma_\nu(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_{\nu+1})\, d\boldsymbol{r}_1 \ldots d\boldsymbol{r}_{\nu+1}, \tag{4.4.6}$$

so that the integral of the sum of products with the same $\nu + 1$ vertices together in irreducible functions γ_ν is

$$V \prod_{\nu \geqslant 1} (\nu!\beta_\nu)^{m_\nu}.$$

One may then write, using (4.3.6) for b_n, that

$$b_n = \sum_{m_1, \ldots, m_\nu \geqslant 0\ (\sum_\nu \nu m_\nu = n-1)} \frac{C(m_1, \ldots, m_\nu, \ldots)}{n!} \prod_{\nu \geqslant 1} (\nu!\beta_\nu)^{m_\nu}, \tag{4.4.7}$$

where $C(m, \ldots, m_\nu, \ldots)$ is the total number of ways that we can arrange the n numbered vertices in the irreducible functions, according to the prescription which permits only a single path of connections through nodal vertices between any pair of irreducible functions.

We now undertake to prove that

$$C(m_1, \ldots, m_\nu, \ldots) = \frac{n!}{\Pi(\nu!)^{m_\nu} m_\nu!}\, n^{m-2}, \tag{4.4.8}$$

so that

$$n^2 b_n = \sum_{m_1, \ldots, m_\nu \geqslant 0\ (\Sigma \nu m_\nu = n-1)} \prod_{\nu \geqslant 1} \frac{(n\beta_\nu)^{m_\nu}}{m_\nu!}. \tag{4.4.9}$$

Consider the complete product of m_1 irreducible functions, γ_1, of two numbered vertices, . . ., m_ν functions γ_ν of $\nu + 1$ numbered vertices. Each function has at least one vertex of order unity or greater. "Erase" the designation n on the last of the numbered vertices, thus leaving an unnumbered vertex, which we shall call a "vacancy", in at least one function if $\lambda_n = 0$, and in $\lambda_n + 1$ functions in the general case. Now erase the designation in λ_α functions of every vertex α of order $\lambda_\alpha > 0$, but in such a way as to leave exactly one vacancy only in every function. This can be done in one and only one way.

Thus every fixed product in which every function γ_ν has a definite assignment of $\nu + 1$ numbered vertices, $1 \leqslant i \leqslant n$, corresponds to one and only one assignment of the first $n - 1$ vertices, ν numbered vertices to each function γ_ν, which can be done in

$$(n - 1)!/\prod_{\nu \geqslant 1} (\nu!)^{m_\nu} m_\nu! \text{ ways.} \tag{A}$$

If we now prove that after this assignment the number of ways that the nodal vertices can be selected and assigned to the functions in an acceptable manner, namely so that one vertex never occurs twice in one function, is

$$n^{m-1} \text{ ways} \tag{B}$$

then we will have proved (8) as (A) × (B).

If the order, λ_α, of each vertex α for $1 \leqslant \alpha \leqslant n$ is given we can assign the vertices of order λ_α greater than unity to the vacancies in the functions in

$$(m - 1)!/\prod_\alpha \lambda_\alpha! \text{ ways.} \tag{C}$$

This can be seen as follows. Select any vertex other than the last vertex n, say vertex α of $\lambda_\alpha > 1$. Enter the designation α to the "vacancies" in λ_α of any of the $m - 1$ functions other than that in which α already occurs. This can be done in $(m - 1)!/\lambda_\alpha (m - 1 - \lambda_\alpha)!$ ways. We now have $m - \lambda_\alpha$ functions with "vacancies". For the next vertex β with $\lambda_\beta > 1$ the number of ways it can be assigned to vacancies in functions other than that in which it already occurs is $(m - 1 - \lambda_\alpha)!/\lambda_\beta!(m - 1 - \lambda_\alpha - \lambda_\beta)!$ ways. Taking the continued product, finally assigning the last unlocated vertex n to the remaining $\lambda_n + 1$ vacancies in only one unique way, we obtain (C). But (C) is the coefficient of $\prod_\alpha x_\alpha^{\lambda_\alpha}$ in

$$(\sum_{\alpha=1}^{\alpha=n} x_\alpha)^{m-1}.$$

The sum of these coefficients over all values of λ_α, $\Sigma\lambda_\alpha = m - 1$, is n^{m-1} or the expression given under (B). We have therefore proven (8) and (9).

With this general expression (9) for n^2b_n, correct for all n values, we now turn to the general proof of (4.3.7) and (4.3.12) giving P/kT and z in terms of ρ. Equation (9) for n^2b_n shows that it is the coefficient of y^{n-1} in the expansion of the generating function $\exp n\Sigma\beta_\nu y^\nu$. By Cauchy's theorem this can be written as a contour integral of y around the origin,

$$n^2b_n = \frac{1}{2\pi i}\oint y^{-n}[\exp n \sum_{\nu>1} \beta_\nu y^\nu]dy, \tag{4.4.10}$$

where the path of integration must enclose the origin but no other singularity of the integrand.

Multiply (10) by z^n on both sides. The integrand is then the nth power of $(x/y)\exp\Sigma\beta_\nu y$. Sum over n from unity to infinity, changing the order of summation and integration to write

$$\sum_{n\geqslant 1} nb_nz^n = \frac{1}{2\pi i}\oint dy \sum_{n\geqslant 1}\left[\frac{z}{y}\exp\sum_{\nu-1}\beta_\nu y^\nu\right]^n. \tag{4.4.11}$$

For sufficiently small z values we can choose the path of integration with $|y| > y_0$ where

$$y_0 \exp - \sum_{\nu\geqslant 1}\beta_\nu y_0^\nu = z, \tag{4.4.12}$$

in which case, since the integrand term is less than unity,

$$\sum_{n\geqslant 1} x^n = x/(1-x), \quad x < 1,$$

we have

$$\sum_{n-1} n^2bz^n = \frac{1}{2\pi i}\oint dy \frac{z}{y[\exp - \Sigma\beta_\nu y^\nu] - z}. \tag{4.4.13}$$

The path of integration now encloses the pole at $y = y_0$ of equation (12). The residue at this pole is obtained by replacing the denominator with its derivative. The derivative is

$$\begin{aligned}(d/dy)[y\exp - \sum_{\nu\geqslant 1}\beta_\nu y^\nu]_{y=y_0} &= [1 - \sum_{\nu\geqslant 1}\nu\beta_\nu y_0^\nu]\exp - \sum_{\nu\geqslant 1}\beta_\nu y_0^\nu \\ &= [1 - \sum_{\nu\geqslant 1}\nu\beta_\nu y_0^\nu](z/y_0),\end{aligned} \tag{4.4.14}$$

from (12). One finds,

$$\sum_{\nu\geqslant 1} n^2 b_n z^n = y_0[1 - \sum_{n\geqslant 1} \nu\beta_\nu y_0^\nu]^{-1}, \tag{4.4.15}$$

which is correct for small values of z.

Now from (12)

$$d \ln z/dy_0 = y_0^{-1}[1 - \sum_{\nu\geqslant 1} \nu\beta_\nu y_0^\nu], \tag{4.4.16}$$

and (15) becomes

$$dy_0 = \sum_{n\geqslant 1} n^2 b_n z^n \, d \ln z. \tag{4.4.17}$$

This integrates to

$$y_0 = \sum_{n\geqslant 1} n b_n z^n, \tag{4.4.18}$$

and comparison with (4.3.9) that $\rho = \Sigma n b_n z^n$ identifies y_0 with the number density, ρ. Finally, using (16) again to integrate $y_0 d \ln z$,

$$\int y_0 \, d \ln z = \sum_{n\geqslant 1} b_n \Sigma^n = y_0 \left[1 - \sum_{\nu\geqslant 1} \frac{\nu}{\nu + 1} \beta_\nu y_0^\nu\right]. \tag{4.4.19}$$

We now have, from (12), and the identification of y_0 with ρ, that

$$\ln z = \ln \rho - \sum_{\nu\geqslant 1} \beta_\nu \rho^\nu, \tag{4.4.20}$$

whereas (19) with (4.2.12) gives

$$P/kT = \sum_{N\geqslant 1} b_n z^n = \rho\left[1 - \sum_{n\geqslant 1} \frac{\nu}{\nu + 1} \beta_\nu \rho^\nu\right]. \tag{4.4.21}$$

Differentiation of (21) with respect to ρ gives the isothermal compressibility κ,

$$\kappa^{-1} = -V^{-1}(\partial P/\partial V)_T = \rho(\partial P/\partial \rho)_T, \tag{4.4.22}$$

$$\kappa^{-1} = \rho kT[1 - \sum_{\nu\geqslant 1} \nu\beta_\nu \rho^\nu]. \tag{4.4.23}$$

From (15) this is also

$$\begin{aligned} \kappa^{-1} &= \rho^2 kT / \sum_{n\geqslant 1} n^2 b_n z^n \\ &= kT(\sum_{n\geqslant 1} n b_n z^n)^2 / \sum_{n\geqslant 1} n^2 b_n z^n. \end{aligned} \tag{4.4.24}$$

These equations are valid for sufficiently small values of z or ρ. The contour integration of (11) was restricted to enclose the origin of

y and no other singularity. On the other hand to use (13) it was necessary that it be taken along a path $|y| > y_0 = \rho$. This step, then is valid as long as the sum $\Sigma\beta_\nu\rho^\nu$ converges. However, from (15) $\Sigma n^2 b_n z^n$ is singular if $\Sigma\nu\beta_\nu\rho^\nu = 1$. We have, therefore, that the equations above are valid as long as ρ is smaller than both ρ_s and ρ_1,

$$\rho < \rho_s, \quad \rho < \rho_1, \tag{4.4.25}$$

where ρ_s is the smallest value of ρ for which

$$\sum_{\nu\geqslant 1} \beta_\nu \rho_s^{\,\nu} = \text{singular} \tag{4.4.26}$$

and ρ_1 is the (first) value of ρ for which

$$\sum_{\nu\geqslant 1} \nu\beta_\nu \rho_1^{\,\nu} = 1. \tag{4.4.27}$$

We shall later associate the values $\rho = \rho_1$ or ρ_s (and actually always ρ_s is smaller than ρ_1 below the critical point) with the value of the number density ρ at condensation.

4.5. Interpretation of $P/kT = \Sigma b_n z^n$

There exists a very simple interpretation of the equation for PV, namely that we can define average numbers, $\overline{N}_1$, $\overline{N}_2$, . . ., $\overline{N}_n$, of monomers, dimers, . . ., n-mers, in the gas, such that

$$PV = kT \sum_{n\geqslant 1} \overline{N}_n. \tag{4.5.1}$$

Consider first the general problem of deciding how many associated complexes of different numbers of points we have if a fixed spatial array of N points is given in a three-dimensional space. Obviously if the average distance, $\rho^{-1/3} = (V/N)^{1/3}$, of the points is very large, and a considerable fraction of the points were paired at a distance $r_0 << \rho^{-1/3}$, we could describe the arrangement as consisting of single points plus a fraction of them paired as dimers. But should we count all pairs whose distance, r_{ij}, is less than $\rho^{-1/3}$ as dimers? Obviously not, since a purely random distribution would give a considerable fraction of the pairs at distances less than $\rho^{-1/3}$. However, we could compute, from the knowledge of ρ alone, the average number of pair distances between r and $r + dr$ if the distribution were purely random. The actual number minus the purely random number, integrated for all distances would converge if there were no long range order, and this number, $\overline{N}_2$ would be the *excess* of associated pairs over that expected for a random distribution.

In a similar way we could define the number, $\bar{N}_3$, of triples in excess of that expected from a random distribution, and in general, $\bar{N}_n$, for n-mers. Since the total number, N, of points is fixed, we would like to require that

$$\sum_{n\geqslant 1} n\bar{N}_n = N. \qquad (4.5.2)$$

In general the numbers, $\bar{N}_n$, may, however, be negative. The distribution may be such that fewer pairs are found close together than randomly expected, and similarly fewer n-mers. In this case we may well have $N_1 > N$, and in general $\Sigma N_n > N$ is possible.

Now the probability density of the distribution of the N coordinates in V for N molecules in V is proportional to exp $-U/kT$. In writing (4.2.4) as a sum of products of f_{ij}'s, which we later summed and associated in a sum of products of cluster functions, the individual cluster function g_n is exactly proportional to the excess probability density (positive or negative depending on the value of the coordinates) of finding a cluster of n molecules at the coordinate position of the configuration space. The integral b_n is then proportional to the excess probability of finding this cluster close together. The proportionality factor is such that,

$$\bar{N}_n = Vb_n z^n, \qquad (4.5.3)$$

as can be seen from eq. (4.3.9) for ρ,

$$\rho V = \bar{N} = \sum_{n\geqslant 1} nVb_n z^n = \sum_{n\geqslant 1} n\bar{N}_n, \qquad (4.5.4)$$

which agrees with (2). Using (3) in (4.2.12) for $PV/kT = V\Sigma b_n z^n$ one finds eq. (1).

Thus by defining the quantities $\bar{N}_n$ as the average number of clusters of molecules into n-mers in excess of the purely random distribution one may say that PV/kT is always equal to the total number of clusters. The statement, however, completely loses meaning in the condensed phase.

4.6. Evaluation of the virial coefficients

The term $-[\nu/(\nu + 1)]\beta_\nu$ which is the coefficient of $\rho^{\nu+1}$ in the expansion (4.4.21) of P/kT as a power series in the number density ρ, is called the $(\nu + 1)$th virial coefficient. The first virial coefficient is unity. Most frequently the macroscopic virial coefficient, $-N_0^\nu[\nu/(\nu + 1)]\beta_\nu$, which is the coefficient of $V^{-\nu}$ in the expansion of PV/RT with V the mol volume, is given in the literature. The integrals,

β_ν, for $\nu = 1$ and 2 have been evaluated theoretically from (4.4.1) and (4.4.2) using several different potential functions.† We shall discuss briefly only a very few cases.

The "hard sphere" potential, corresponding to rigid spheres of radius r_0, without attraction, is defined by,

$$\begin{aligned} u(r) &= \infty, \quad 0 \leqslant r \leqslant 2r_0, \\ u(r) &= 0, \quad 2r_0 < r < \infty. \end{aligned} \tag{4.6.1}$$

The volume of the individual sphere is then

$$v_0 = \frac{4\pi}{3} r_0^3. \tag{4.6.2}$$

The function $f(r_{ij})$ is -1 for $r \leqslant 2r_0$ and zero for $2r_0 < r$, so that, from (4.3.2)

$$\beta_1 = -8v_0, \tag{4.6.3}$$

and the second virial coefficient is

$$B_2 = \tfrac{1}{2}N_0\beta_1 = 4N_0v_0, \tag{4.6.4}$$

namely 4 times the sum of the volumes of the hard spheres. For this hard sphere potential the higher virial coefficients up to the fifth are

$$\begin{aligned} B_3 &= \frac{5}{8} B_2^2 = -\frac{2}{3} N_0^2\beta_2, & \beta_2 &= -15v_0^2, \\ B_4 &= 0.2869\, B_2^3 = -\frac{3}{4} N_0^3\beta_3, & \beta_3 &= -24.48v_0^3, \quad (4.6.5) \\ B_5 &= (0.115 \pm 0.005)B_2^4 = -\frac{4}{5} N_0^4\beta_4, & \beta_4 &= -37v_0^4. \end{aligned}$$

If there is a "square well" attractive potential in addition to the hard sphere repulsion,

$$\begin{aligned} u(r) &= \infty, & 0 &\leqslant r \leqslant 2r_0, \\ &= -u_0, & 2r_0 &< r \leqslant 2r_0 + \Delta r, \\ &= 0, & 2r_0 + \Delta r &< r < \infty, \end{aligned} \tag{4.6.6}$$

---

† A good discussion, with extensive tables is given in *Molecular Theory of Gases and Liquids*, Hirschfelder, Curtiss and Bird, John Wiley and Sons, N.Y., Chapman and Hall, London, 1954, but a very considerable number of later computations have been made.

then

$$\beta_1 = -4v_0 + 4v_0\left\{\left[1 + \frac{\Delta r}{2r_0}\right]^3 - 1\right\}[e^{u_0/kT} - 1].$$

At sufficiently high temperatures $e^{u_0/kT} - 1 = (u_0/kT)$ and the extra term is inversely proportional to the temperature. This is the form of the development of the van der Waals equation.

The most generally used realistic potential is the Lennard-Jones 6–12 potential having a minimum energy of $-\varepsilon_0$ at $r = r_0$, rising rapidly to a zero value at $r = r^*$,

$$r_0 = 2^{1/6}\, r^* = 1.220r^*,$$

$$u(r) = \varepsilon_0\left[\left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^6\right] = 4\varepsilon_0\left[\left(\frac{r^*}{r}\right)^{12} - \left(\frac{r^*}{r}\right)^6\right]. \qquad (4.6.7)$$

This is a two-parameter expression and, with $v_0 = (\pi/6)r_0^3$ or $v^* = (\pi/6)(r^*)^3$ the $(\nu + 1)$th virial coefficient, β_ν, divided by v_0^ν or $(v^*)^\nu$ depends only on (ε_0/kT). Rather extensive tables exist for this potential.

There is considerable evidence that the actual potential cannot be given with any high degree of precision by a pure sum of pair potentials, and that, rather than (4.2.1) the form

$$U_N = \sum_{N \geqslant i > j \geqslant 1}\sum u_2(r_{ij}) + \sum_{N \geqslant i > j > k \geqslant 1}\sum\sum u_3(r_{ij}, r_{jk}, r_{ki}) + \Sigma\Sigma\Sigma\Sigma\, u_4 + \cdots \qquad (4.6.8)$$

would be necessary to give the true potential energy function with accuracy. There is no difficulty in incorporating the term u_3 into the third virial coefficient expression using the cluster function technique, and u_4 into the fourth, etc., although needless to say the integrals would become more painful to evaluate. So little, however, is known about the true nature of the terms in (8) that the present art contents itself with (4.2.1), usually using the Lennard-Jones 6–12 potential of eq. (7).

In section 4.8 we discuss the one case in which we actually believe that the exact potential is known, at least for large values of r, and that it is really the sum of pair terms of (4.2.1), but in which the decay to zero of $u(r)$ with large r is so slow that the individual cluster integrals diverge.

4.7. Condensation

The thermodynamic description of condensation is that the density, at constant temperature, goes from a value ρ_g for the saturated vapor, to some higher value, ρ_1 for the condensed phase, at constant value of P and activity z. Since $[\partial(P/kT)/\partial \ln z]_T = \rho$ this means that there is an abrupt change of slope in the curve of P/kT versus z. By definition a discontinuity of derivative in a function occurs only at a singularity of that function.

Now the original expression (4.1.6) for exp PV/kT gives it as a power series in z^N with coefficient $Q_\tau(N,V,T)/N!$. If the potential energy, U_N, is given by the pair sum of (4.2.1) and $u(r_{ij}) \rightarrow \infty$ for r less than some r_0, in short if there is a true hard sphere repulsion at r_0, then for any finite V there will be some maximum number, $N_{\max}$, of molecules which can fit in V at finite potential energy. For $N > N_{\max}$ the integrand, $\exp - U_N/kT$ of $Q_\tau(N,V,T)$ is everywhere zero, $Q_\tau = 0$, and the sum of powers of z terminates at $N_{\max}$. We have that $\exp(PV/kT)$ is given by a polynomial in z of order $N_{\max}$. The real case cannot be greatly different.

It is a fundamental theorem of mathematics that a polynomial does not have a singularity for finite values of its argument, so that P/kT must be an analytic function of z and can have no discontinuity in derivative.

The semantic difficulty is removed by a simple artifice. Let the volume V become infinite and write

$$P/kT = \lim_{V\to\infty} \left[V^{-1} \ln \sum_{N\geqslant 0} \frac{z^N}{N!} Q_\tau(N,V,T) \right]. \tag{4.7.1}$$

For infinite V the sum runs to infinite N. It is no longer a polynomial of finite order but an infinite series, and singularities are now permitted.

The difficulty and its resolution have led to the somewhat curious statement upon occasions that statistical mechanics is valid only for systems of infinite volume. It would appear to be more realistic to say that the condition for equilibrium between two thermodynamic phases permits any fraction between zero and one of the material in one phase for pressures or activities that do not differ within experimental error, and thus forego the luxury of the more elegant mathematical formulation of a singularity in the $P(z,T)$ function.

Actually, of course, there is no difficulty whatsoever in constructing a polynomial of order 10^{20} with a change of slope sufficiently abrupt that any measurement would make it appear discontinuous. Even the most simple two-term expression,

$$P/\rho kT = N^{-1}\ln[a^N z^N + a^N b^{-N} z^{2N}], \tag{4.7.2}$$

with $a > 0$, $b > 0$, and, say, $N = 10^{20}$ shows a far more abrupt change of slope at $z = b$ than is necessary to justify this statement. For $z = b(1 - \varepsilon)$ even at $\varepsilon = 10^{-20}$ we can write

$$\begin{aligned} P/\rho kT &= [\ln az] + N^{-1}\ln[1 + (1-\varepsilon)^N] \\ &= \ln ab + \ln(z/b) + O(1/N), \end{aligned} \tag{4.7.3}$$

which has the slope unity to within one part in 10^{20}, since, for the numbers given $(1 - \varepsilon)^N = e^{-1}$. For $z = b(1 + \varepsilon)$ and again $\varepsilon = 10^{-20}$, we find

$$P = \ln ab + 2\ln(z/b) + O(1/N) \tag{4.7.4}$$

so that the slope doubles in relative change of z by one part in 10^{20}.

However, although it is hardly realistic or necessary to explain the very abrupt change in properties with change of thermodynamic phase, the formulation of eq. (1) does have a considerable usefulness, since it permits us to identify condensation with the first singularity on the positive z-axis of the function $P(z)$ defined by (1) at infinite volume.

With the definition (4.2.6) of $b_n(V)$, but as yet not defining the limit b_n of $b_n(V)$ at $V \to \infty$, eq. (4.2.7), we can write eq. (4.2.12) rigorously as

$$P(V,z,T) = kT\sum_{n\geqslant 1} b_n(V)z^n, \tag{4.7.5}$$

and, for the limit

$$P = \lim_{V\to\infty}[P(V,z,T)], \tag{4.7.6}$$

$$P/kT = \lim_{V\to\infty}[\sum_{n\geqslant 1} b_n(V)z^n]. \tag{4.7.7}$$

As long as the series

$$P/kT = \sum_{n-1}\lim_{V\to\infty}\{[b_n(V)]\}z^n \tag{4.7.8}$$

converges it can be used to replace the limit of the sum, and equation (4.2.12) with the definition (4.2.7) of b_n as the limit of $b_n(V)$ is rigorously correct. For z-values below that of the singularity of the series in the right of (8) the further steps to (4.4.21) are valid. The first singularity of (8) on the positive z-axis is then given by (4.4.25) to (4.4.27).

It does not follow that the first singularities of the two expressions (7) and (8) are identical. Series can be found for which they are different, indeed for certain (very unrealistic) models of the potential for which computations can be made the two are different.

Since for actual potentials the calculations of

$$\lim_{n\to\infty} [(b_n)^{1/n}] = b_0 \text{ or } \lim_{\nu\to\infty} [(\beta_\nu)^{1/\nu}] = \beta_0 \tag{4.7.9}$$

which determine the values of z_s and $\rho_s = b_0^{-1}$, ρ_s, $z_s = \beta_0^{-1}$ for which (4.2.12) and (4.4.21) become singular the question might seem to be of only abstract interest.

There is, however, compelling reason for accepting the formulation (8) at low temperatures. At sufficiently low values of T for which, at the distance r_0 of minimum pair potential, $\exp - [u(r_0)/kT]$ approaches extremely large values, the contribution to b_0 of (9) will come only from configurations of the molecules with a maximum number of pair distances very close to r_0, in short from configurations of small displacements from the crystalline lattice configuration. Since the cluster function g_n includes all connected products we can replace $f(r_{ij}) = [(\exp - (u/kT) - 1]$ for all non-neighboring pairs by $f(r_{ij}) + 1 = \exp - u(r_{ij})/kT$, since this replacement does not introduce disconnected products. For the close neighbors with $\exp - u(r_0)/kT >> 1$ we can approximate $f(r_{ij}) \to f(r_{ij}) + 1 = \exp - u(r_{ij})/kT$, but must limit the integration to that part of configuration space for which all molecules remain within range of each other's forces, thus removing the part proportional to V^N in $Q_\tau(N,V,T)$. But this replacement means just that we replace $b_N(V)$ by $V^{-1}Q_\tau(N,V,T)$ restricted to the configurations for which the molecules form a crystal. In turn this will make b_0 of (9) equal to the reciprocal, z_c^{-1}, of the activity z_{cryst}, of the low temperature crystal. Thus pragmatically (8) does give the correct condensation, $z_s = z_{\text{cryst}}$ in the limit of $T \to 0$, and we may surmise that it does also at higher temperatures.

At sufficiently high temperatures, namely above critical, there is presumably no singularity in the $P(z)$ curve. The critical temperature is that for which the first singularity, ρ_s of (4.4.26), crosses the value, ρ_1, for which the sum of (4.4.27) is unity, and it appears that above this temperature the singularities of (7) and (8) differ. This is mere speculation, but is based on the fact that if ρ_1 were to determine the density of condensation then we would require, from (4.4.23) and (4.4.27) that the compressibility, κ, of the gas go continuously to an infinite value as the condensation density is approached. This does not happen experimentally for any continuous temperature range, but only at the critical temperature. It is somewhat difficult to believe that for all the mutual potentials that exist the singularity ρ_s of (4.4.26) disappears off the real axis into the complex plane exactly at the temperature and density for which the series of (4.4.27) reaches a maximum value unity, and then descends.

4.8. Plasma and the Debye limit

The case of a gas of charged particles presents us with a mutual pair potential proportional to r^{-1}. At large distances $f_{ij} = e^{-u/kT} - 1 \cong -u/kT$ goes to zero as r^{-1} and $\int 4\pi r^2 f\,dr$ diverges badly. The answer to the difficulty lies in the fact that there are always equal numbers of positively and negatively charged species, so that, although the individual terms approach infinity there are equal positive and negative contributions which cancel. However, it is not only the two-particle terms that diverge, but other terms lead to divergent integrals, and a moderately sophisticated treatment of summation in a given order before integration is necessary.

The result of the treatment is to give a first-order correction proportional to $\rho^{1/2}$ rather than a virial development as a power series in the number density, ρ. This square root correction was first found by Peter Debye in the twenties, as the correction for electrolytic solutions. Since, in Chapter 6 on dense fluids we show that the imperfect gas results can be taken over rigorously to the case of solutions, we will introduce the treatment of a plasma gas of charged particles here as a special case of the cluster development. To many chemists, at least, the greater interest will lie in the application of the equations to solutions, the essential difference being that whereas the gas plasma is in a vacuum of dielectric constant unity, the electrolytic solution is a plasma in a medium of the dielectric constant of the solvent.

The treatment is for a neutral plasma. There must then be at least two species, a and b, of particles, positively and negatively charged, species a having charge $\zeta_a e$ and number density ρ_a. The neutrality condition is then

$$\sum_a \rho_a \zeta_a = 0. \tag{4.8.1}$$

We must therefore extend the treatment of this chapter to a gas of more than one chemical component. We do this first for a general potential. The extension is described in more detail in section 6.9. It is almost trivial up to all equations involving the cluster integrals b_n and activities z, and consists solely of replacing n by $\mathbf{n} = n_a, n_b, \ldots$, $n!$ by $\mathbf{n}! = n_a!, n_b!, \ldots$, z by $\mathbf{z} = z_a, z_b, \ldots$, and z^n by $\mathbf{z}^{\mathbf{n}} = z_a^{n_a}$,

$z_b^{n_b}$, . . . The cluster function $g_{\mathbf{n}}$ is still defined by (4.2.5) but the $f(r_{ij})$'s in the product on the right are now different functions, $f_{aa}(r_{ij}), f_{ab}(r_{ij}), \ldots$, depending on the species, $a, b, \ldots$, to which the molecules i and j belong. Expression (A) in section 4.2 as well as (B) for the combinatorial factor are correct with the replacements $\mathbf{n}$ for n. One arrives at

$$P/kT = \sum_{\mathbf{n}} b_{\mathbf{n}} \mathbf{z}^{\mathbf{n}}, \tag{4.8.2}$$

in which all b's with $n = \Sigma n_a = 1$ have values of unity. We also find

$$\rho_a = \partial(P/kT)/\partial \ln z_a = \sum_{\mathbf{n}} n_a b_{\mathbf{n}} \mathbf{z}^{\mathbf{n}}. \tag{4.8.3}$$

Because of the already involved combinatorial problem of deriving the coefficient $C(m_1, \ldots, m_\nu, \ldots)$ of eq. (4.4.8) the derivation and proof of the development in power series of the number densities in the multicomponent case is very cumbersome. We here present only the results.†

The irreducible functions, $\gamma_\nu(\mathbf{r}_1, \ldots, \mathbf{r}_{\nu+1})$ of eq. (4.4.1) were functions of the coordinates of $\nu + 1$ molecules. We replace $\nu + 1$ by $\mathbf{m} = m_a, m_b, \ldots, m = \Sigma m_a$ and define a function $\Gamma_{\mathbf{m}}(\mathbf{r}_1, \ldots, \mathbf{r}_m)$ of m molecules in the same way as γ_{m-1},

$$\Gamma_{\mathbf{m}}(\mathbf{r}_1, \ldots, \mathbf{r}_m) = \sum_{x_{ij}=0,1} \begin{bmatrix} \text{sum of connected} \\ \text{products without} \\ \text{nodes} \end{bmatrix} \prod_{m \geqslant i > j \geqslant 1} [f_{ab}(i, j)]^{x_{ij}} \tag{4.8.4}$$

Then replace the quantities β_ν of (4.4.6) by

$$B_{\mathbf{m}} = \frac{1}{V\mathbf{m}!} \int\!\!\int_V \cdots \int \Gamma_{\mathbf{m}}\, d\mathbf{r}_1 \ldots d\mathbf{r}_m, \tag{4.8.5}$$

so that if $\mathbf{m}$ has $m_a > 0$ for only one component, then

$$B_{\mathbf{m}} = m^{-1} \beta_{m-1}.$$

One finds instead of (4.4.20),

$$\ln z_a = \ln \rho_a - \rho_a^{-1} \sum_{m \geqslant 2} m_a B_{\mathbf{m}} \boldsymbol{\rho}^{\mathbf{m}}, \tag{4.8.6}$$

---

† The proofs are given by J. E. Mayer, *J. Phys. Chem.*, **43**, 71 (1939) for two components and by K. Fuchs, *Proc. Roy. Soc.* **A 179**, 408 (1942) for the general case.

for the activity, z_a, of each component a. It is convenient, for our purpose, to define the sum

$$S = \sum_{m>2} B_m \rho^m \tag{4.8.7}$$

so that (6) can be written as

$$\ln (z_a/\rho_a) = -\partial S/\partial \rho_a. \tag{4.8.8}$$

With ρ for the total number density,

$$\rho = \sum_a \rho_a,$$

eq. (4.4.21) for the pressure becomes

$$\begin{aligned} P/kT &= \rho - \sum_{m \geqslant 2} (m-1) B_m \rho^m \\ &= \rho + S - \sum_a \rho_a \, \partial S/\partial \rho_a. \end{aligned} \tag{4.8.9}$$

Now these results are valid only for molecules with short range mutual potentials that fall off with the distance r more rapidly than r^{-3}. To evaluate the equations for the r^{-1} electrostatic potential we consider first a mutual pair potential of the form $r^{-1} e^{-\alpha r}$ which, for positive $\alpha > 0$ satisfies the condition of leading to convergent integrals and coefficients b_n, B_m, which have finite limiting values as $V \to \infty$. We then later go to the asymptotic limit at $\alpha \to 0$.

In the first place we examine the methods of obtaining certain integrals which we need later. Define I_0 and I_n for $n \geqslant 1$,

$$\begin{aligned} I_0(r_{ij}) &= f(r_{ij}), \\ I_n(r_{ij}) &= \int f(r_{i1}) f(r_{12}) \ldots f(r_{n-1,n}) f(r_{nj}) d\mathbf{r}_1 \ldots d\mathbf{r}_n. \end{aligned} \tag{4.8.10}$$

The integrals can be evaluated by "folding." One introduces $g(t)$ as the three-dimensional Fourier transform of $f(r)$,

$$g(t) = \int_0^\infty 4\pi r^2 f(r) \frac{\sin 2\pi t r}{2\pi t r} \, dr. \tag{4.8.11}$$

Then

$$f(r) = \int_0^\infty 4\pi t^2 g(t) \frac{\sin 2\pi t r}{2\pi t r} \, dt, \tag{4.8.12}$$

and[†]

$$I_n(r) = \int_0^\infty 4\pi t^2 [g(t)]^{n+1} \frac{\sin 2\pi t r}{2\pi t r} \, dr. \tag{4.8.13}$$

[†] The method is described in various texts, for this and similar applications see E. W. Montroll and J. E. Mayer, *J. Chem. Phys.*, **9**, 626 (1941).

In the particular case that $r = 0$, $(\sin 2\pi tr/2\pi tr) \to 1$ and

$$I_n(0) = \int_0^\infty 4\pi t^2[g(t)]^{n+1}\,dt. \tag{4.8.14}$$

Define

$$J(\kappa,r) = \sum_{n>0} (-\kappa^2)^n I_n(r). \tag{4.8.15}$$

From (13) with

$$G(\kappa,t) = g(t) \sum_{n>0} [-\kappa^2 g(t)]^n, \tag{4.8.16}$$

we find that

$$J(\kappa,r) = \int_0^\infty 4\pi t^2\, G(\kappa,t)\, \frac{\sin 2\pi tr}{2\pi tr}\, dr, \tag{4.8.17}$$

and at $r = 0$,

$$J(\kappa,r = 0) = \int_0^\infty 4\pi t^2\, G(\kappa,t) dt. \tag{4.8.18}$$

As long as $g(t)$ has a finite upper bound we can always choose κ^2 sufficiently small so that $\kappa^2 g(t) < 1$ for all values of t, in which case

$$G(\kappa,t) = g(t)[1 + \kappa^2 g(t)]^{-1} = \{\kappa^2 + [1/g(t)]\}^{-1},\ (\kappa^2 g < 1). \tag{4.8.19}$$

We now investigate the hypothetical case that

$$f(r) = (4\pi r)^{-1}\, e^{-\alpha r}, \tag{4.8.20}$$

for which, from (11) we find,

$$g_\alpha(t) = [\alpha^2 + (2\pi\, t)^2]^{-1}.$$

As long as $0 < \kappa < \alpha$ we can use (19) for $G_\alpha(\kappa,t)$ finding,

$$G_\alpha(\kappa,t) = [\kappa^2 + \alpha^2 + (2\pi\, t)^2]^{-1}, \tag{4.8.21}$$

so that, from (17),

$$J_\alpha(\kappa,r) = (4\pi r)^{-1} \exp - \sqrt{(\kappa^2 + \alpha^2)}r. \tag{4.8.22}$$

We shall later use the analytical continuation of (22) as we let $\alpha \to 0$ to write

$$J_0(\kappa,r) = (4\pi r)^{-1} \exp - \kappa r. \tag{4.8.23}$$

Those familiar with the Debye theory will recognize this as the analytical form of the potential of average force between two ions in the Debye limit with κ^{-1} the "Debye length." Its value is

$$\begin{aligned}\kappa^2 &= (4\pi e^2/DkT)\sum_a \rho_a \zeta_a{}^2 \\ &= (4\pi e^2/DkT)\rho\sum_a x_a\zeta_a{}^2 \qquad (4.8.24)\end{aligned}$$

for ions in a solvent of dielectric constant D having mol fraction $x_a, x_b, \ldots, \Sigma x_a\zeta_a = 0$.

We have discussed the necessary equations. We now examine the case of a dilute gas plasma, in which case $D \equiv 1$. The mutual potential between two ions has a short range, largely repulsive, part and the long range electrostatic expression $(\zeta_a\zeta_b e^2/r)$. Assume the short range potential to become identically zero for $r > r_0$ and define

$$\begin{aligned}f_{ab}{}^* &= f_{ab}(r) = [\exp - (u_{ab}(r)/kT)] - 1, \quad 0 \leqslant r \leqslant r_0, \\ f_{ab}{}^* &= 0 \qquad\qquad r_0 < r \leqslant \infty. \qquad (4.8.25)\end{aligned}$$

For r greater than r_0 we replace the electrostatic potential by

$$\begin{aligned}&\zeta_a\zeta_b e^2/rkT \rightarrow (4\pi\zeta_a\zeta_b e^2/kT)\phi_\alpha(r) \\ &\phi_\alpha(r) = 0, \quad 0 \leqslant r \leqslant r_0, \\ &\phi_\alpha(r) = (1/4\pi r)e^{-\alpha r}, \quad r_0 < r \leqslant \infty, \qquad (4.8.26)\end{aligned}$$

so that upon expanding the exponential in f_{ab} for $r > r_0$ we have

$$f_{ab}(r) = f_{ab}{}^*(r) + \sum_{n\geqslant 1}\left[-\frac{4\pi\zeta_a\zeta_b e^2}{kT}\right]^n \frac{1}{n!}[\phi_\alpha(r)]^n \qquad (4.8.27)$$

which goes over correctly to the plasma case in the limit $\alpha \rightarrow 0$.

With the sum (27) for f_{ab} in the expression on the right of (4) for Γ_m we now sum products of functions in which for every ij pair we may have unity, $f_{ab}{}^*$, or some power of ϕ_α. Using a distinctive line for $f_{ab}{}^*$ we can, as before, make a one-to-one correspondence between products and graphs having no nodal vertices, but with n different ϕ_α-lines between the same two vertices, ij, to represent $\phi_\alpha{}^n$.

We now proceed to order the graphs according to types, which we shall define, and can then consider the total contribution of each type, in order, to the sum S of eq. (7).

The first type, the linear type, occurs only for $m = 2$ and is that with one ϕ_α-bond between the two molecules corresponding to the term $n = 1$ in the sum (27). The function is $-(4\pi\zeta_a\zeta_b e^2/kT)\phi_\alpha(r_{ij})$. The contribution to S is obtained by first multiplication with $(\rho_a\rho_b/\mathbf{m}\,!)$ and summation over a, and subsequent integration of the function. Before integration we have

$$-(4\pi e^2/kT)\phi_\alpha(r)(\sum_a \tfrac{1}{2}\rho_a{}^2\zeta_a{}^2 + \sum_{a \neq b}\sum \rho_a\zeta_a\rho_b\zeta_b)$$

$$= -(4\pi e^2 kT)\phi_\alpha(r)\tfrac{1}{2}[\sum_a \rho_a\zeta_a]^2 = 0 \qquad (4.8.28)$$

from the neutrality condition, eq. (1). The contribution to S is zero.

The second type are the cycles, contributing $S^{(c)}$ to S, and it is these which give corrections proportional to $\rho^{1/2}$ to the thermodynamic functions. The cycle graph for $m = 2$ consists of those having $\phi_\alpha{}^2$ in the sum (27), but for $m > 2$ they are those having a cycle of $m\phi_\alpha$-bonds, $\phi_\alpha(r_{12})\phi_\alpha(r_{23}) \ldots \phi_\alpha(r_{m1})$. Neglecting the contribution to the integral at the range any $r_{ij} \leqslant r_0$ which can, for these graphs, be absorbed in the short range function $f_g{}^*$, the integral of any such cycle is the value if $I_{n-1}(r{=}0)$ defined by (10) with $f = \phi = (4\pi r)^{-1}e^{-\alpha r}$.

Since each particle i contributes to two ϕ-bonds there is a factor

$$\prod_a (-4\pi\rho_a\zeta_a{}^2e^2/kT)^{m_a}$$

multiplying the cycle product of ϕ_α's for each completely numbered graph. This must be divided by $\mathbf{m}\,!$ to give its contribution to $B_{\mathbf{m}}$. For $m > 2$ it must now be multiplied by $\frac{1}{2}(m-1)\,!$ which is the number of ways that m numbered objects can be arranged in a ring. For $m = 2$ it must be multiplied by the factor $(1/2\,!)$ of $\phi_\alpha{}^2$ in the sum (27) but $\frac{1}{2} = \frac{1}{2}(m-1)$ if $m = 2$ so that $\frac{1}{2}(m-1)\,!$ is general. We then sum over all $\mathbf{m}$ for fixed m, finding, from (24) for κ^2 with $D = 1$,

$$\frac{1}{2m}\left\{\sum_{m_a} \frac{m\,!}{\prod_a m_a\,!}\left[\frac{-4\pi\rho_a\zeta_a{}^2e^2}{kT}\right]^{m_a}\right\}_{\Sigma m_a = m} = \frac{1}{2m}\,(-\kappa^2)^m. \qquad (4.8.29)$$

This then is the summed factor to multiply each cycle of $\phi(r_{ij})$'s of fixed m's to obtain its contribution to $S^{(c)}$ We have

$$S^{(c)} = \frac{1}{2}\sum_{n \geqslant 1}(-\kappa^2)^{n+1}\,\frac{1}{n+1}\,I_n(\alpha, r = 0). \qquad (4.8.30)$$

Differentiate with respect to κ^2 to find

$$(\partial S^{(c)}/\partial\kappa^2) = -\tfrac{1}{2}\sum_{n > 1}(-\kappa^2)^n\,I_n\,(\alpha, r = 0)$$

$$= \tfrac{1}{2}[I_0(\alpha, r = 0) - J_\alpha(\kappa, r = 0). \qquad (4.8.31)$$

Finally with (10), $I_0 = \phi_\alpha(r)$, with (26) for $\phi_\alpha(r)$ and (22) for $J_\alpha(\kappa,r)$

$$(\partial S^{(c)}/\partial\kappa^2) = \lim_{\substack{\alpha\to 0\\ r\to 0}} \frac{1}{8\pi} r^{-1}[e^{-\alpha r} - e^{-\sqrt{(\kappa^2+\alpha^2)}r}]$$

$$= \lim_{\alpha\to 0} \frac{1}{8\pi}[\sqrt{(\kappa^2+\alpha^2)} - \alpha]$$

$$= \kappa/8\pi, \tag{4.8.32}$$

and, on integration from $\kappa^2 = 0$ to κ^2

$$S^{(c)} = (\kappa^3/12\pi) \tag{4.8.33}$$

The more complicated graph types lead to terms that depend on higher powers of the density.† We note that, from (24)

$$\partial\kappa^2/\partial\rho_a = (4\pi e^2/kT)\zeta_a^2,$$

$$\sum_a \rho_a(\partial\kappa^2/\partial\rho_a) = \kappa^3. \tag{4.8.34}$$

We may bring out the dependence of κ on ρ explicitly by defining $x_a = \rho_a/\rho$,

$$\lambda^2 = (4\pi e^2/kT)\Sigma x_a\zeta_a^2, \tag{4.8.35}$$

$$\kappa = \lambda\rho^{1/2}. \tag{4.8.36}$$

From (8) with (32) we have for the activity coefficient γ_a,

$$\ln\gamma_a = \ln(z_a/\rho_a) = -\tfrac{1}{2}(e^2/kT)\zeta_a\lambda\rho^{1/2} + \ldots \tag{4.8.37}$$

and from (9) with (32), (33) and (34)

$$P/kT = \rho(1 - (\lambda^3/24\pi)\rho^{1/2} + \ldots). \tag{4.8.38}$$

We will not here discuss the more complicated graphs in any detail. For cycles in which one or more bonds are of type $f_{ab}{}^*(r)$ we add the negative of $(4\rho\zeta_a\zeta_b e^2/kT)\phi_\alpha(r)$ for $r < r_0$ to each $f_{ab}{}^*(r)$ term in the cycle. This corrects for the erroneous inclusion of the low r-values in the integration for $S^{(c)}$. Every such short range term $f_{ab}{}^*$ in cycles contributes an extra factor ρ, so that these graphs contribute terms $\rho^{n+1/2}$ with $n \geqslant 1$ to S. The graphs with ϕ^3 between two particles, along with those having three "chains," $J_\alpha(\kappa, r_{ij})$, between particles i and j

---

† The more general types are discussed in J. E. Mayer, *J. Chem. Phys.* **18**, 1426 (1950).

involve integrals of the type $\int_{r_0}^{\infty} r^{-1} e^{-3\kappa r}\, dr$ which contribute a term $\rho \ln \kappa r_0$. These would diverge logarithmically were we to let $r_0 \rightarrow 0$, but eventually lead to a contribution $\rho \ln \rho$ in S. These terms, at small ρ values, are smaller than the $\rho^{1/2}$ term, but larger than the term in ρ due to the two particle $f_{ab}{}^*$ graphs.

CHAPTER 5

SOLIDS

5.1. Overall survey and electronic excitation

Any attempt to write a nearly complete description of the application of statistical mechanics to solid systems would require many volumes much larger than this one. The subject of "Solid State Physics" now comprises two-fifths of the *Physical Review*, and a majority of the articles involve, in some degree or other, statistical mechanical methods. We will discuss only very few salient examples in this chapter.

With reasonable inclusiveness one can characterize all rigid materials into either of two classes, crystals and glasses. The crystals are characterized by having a triply periodic arrangement of molecules in three-dimensional space, and therefore a long range order. It is not essential that the ordering be perfect. A crystal will always have a small number of imperfections even at equilibrium, and there are crystals that have some completely random features.

The glasses are characterized by an absence of long range order, and their quantitative treatment in statistical mechanical terms requires the techniques appropriate to dense fluids, discussed in Chapter 6. Indeed there is no sharp and discontinuous change of properties in going from the high temperature liquid to the glass, and it is customary to regard glasses as supercooled liquids. Included in this category are many non-crystalline but rigid organic materials, especially high polymers, although many of these have structure which have no features of resemblance to either true crystals or the conventional glasses.

Of the crystalline materials one can again, with reasonable sharpness, divide them into metals, insulators, and a technically very important class called semi-conductors. The insulators are those materials for which the lowest excited electronic state has an energy of the order of an electron volt, $\varepsilon/k = 1.1 \times 10^4$ °K, or more, so that, at normal temperatures, no electronic excitation need be considered. The metals are at the other extreme in which there is a continuum of excited electronic states above the ground state, and electronic excitations can occur even at the lowest temperatures. The electrical conductivity of the metal is due to this ease of exciting electrons. The application of an electric field causes transitions into linear combinations of the stationary states which are states of electric flux. The electrical resistance arises from the scattering of the electrons due to the vibrations of the atoms, and increases with increasing temperature.

A semi-conductor is characterized by a considerably lower electrical conductivity than a metal, usually by quite a few orders of magnitude, but also particularly by the fact that the resistance decreases with increasing temperature.

A pure material having no excited electronic states immediately above the ground state, but having a "band" of N states in an energy range starting at ε_0 of the order of a tenth of an electron volt, $\varepsilon/k = 10^3$ °K, above the ground state should be a semi-conductor. The number of excited electrons would be proportional to $\exp - (\varepsilon_0/kT)$ and these excited electrons would conduct, as would the "holes" in the states normally filled at $T = 0$.

However, most semi-conductors do not conform to this model, but arise due to impurities in a material that would otherwise be an insulator.

If two identical atoms, i, j, having electronic orbital levels ψ_i and ψ_j at the same energy ε_0 but located around their respective atomic centers, are brought close together, the electronic states split in energy. The new states, $\psi_b = (1/\sqrt{2})(\psi_i + \psi_j)$ and $\psi_a = (1/\sqrt{2})(\psi_i - \psi_j)$ have energies approximately $\varepsilon_0 - \frac{1}{2}\Delta\varepsilon$ and $\varepsilon_0 + \frac{1}{2}\Delta\varepsilon$ respectively, with $\Delta\varepsilon$ increasing rapidly as the distance, r, between the atoms decreases. If $N = 10^{20}$ such atoms are arranged in a regular crystalline array at a variable distance, r, between nearest neighbors, the N individual atomic levels, ψ_i, $1 \leqslant i \leqslant N$, will split into N orbital levels of differing energies lying in some energy range $\Delta\varepsilon$, depending on r. These levels, then, form a "band" of levels of continuous energy with definite sharp upper and lower energy limits. If the crystalline array is not regular, but the atoms are paired closely, as in Cl_2, Br_2, etc., the band splits with an energy gap in the middle. There will be $\frac{1}{2}N$ orbital levels in the lower band and $\frac{1}{2}N$ in the upper. Due to the electron spin there are two electronic states for each orbital state, so that the regular crystal band has place for $2N$ electrons. If the atoms have only one electron each in the valence state, as do the alkalis, then the band of $2N$ states is only half filled by the N electrons and the crystal is metallic, having a continuum of energy states above the last filled state.

In such cases the energy distribution of the states in the band is approximately that of the translational states of a free electron gas having the volume of the metal. The statistical mechanics of the free electron gas is treated in section 7.4, and this model is taken as the initial approximation in most treatments of the properties of metals.

In other cases, such as the alkaline earths, although there are two valence electrons per atom, the bands due to the rather closely lying *s* and *p* atomic levels overlap in energy, and the free gas approximation, although poorer, is still a useful starting assumption.

The element germanium of column 4 in the Periodic Table forms a typical semi-conductor, when doped with a very small amount of impurity from an element of either column 3, such as gallium, or column 5 such as arsenic. Absolutely pure germanium would be an insulator. In the case of the column 3 element impurity only a relatively small excitation energy is necessary to attach an electron from the "Fermi Sea" of germanium electrons to the unfilled valence state of the impurity. This leaves vacancies at the top of the germanium band, and the vacancies can move as positively charged electrons would. The material is known as *P*-type semi-conductor. The column 5 element impurity, on the other hand, has extra valence electrons which, with relatively small excitation energy, can go into the unfilled first excited band of germanium. In most preparations the number of impurity atoms is made very small, so that the number of conducting electrons or holes is small, even when the energy gaps are not great.

The contribution to the heat capacity due to the electrons is small, even for metals, where the model of the free electron gas is one with a high density of electrons, namely of one or more per atom. This is due to the fact that the energy width, $\Delta\varepsilon$, of the band containing $2N$ electronic states is several electron volts. Relatively few states are in the energy range kT to which most of the electronic excitations are limited, so that the number of excited electrons is small. In eq. (7.4.3) we derive the expression

$$\varepsilon = \frac{3}{5}\varepsilon_0\left[1 + \frac{5}{12}\left(\frac{\pi kT}{\varepsilon_0}\right)^2 - \frac{1}{16}\left(\frac{\pi kT}{\varepsilon_0}\right)^4 + \cdots\right] \tag{5.1.1}$$

for the energy per electron. In eq. (1) the energy quantity ε_0 is the kinetic energy of the highest filled state at $T = 0°\mathrm{K}$ and, in metals, is of the order of electron volts. Differentiation with respect to T gives a heat capacity per electron c_{elec} of

$$c_{\text{elec}} = \pi k \left[\frac{1}{2}\left(\frac{\pi kT}{\varepsilon_0}\right) - \frac{3}{20}\left(\frac{\pi kT}{\varepsilon_0}\right)^3 + \ldots\right]. \tag{5.1.2}$$

At $T = 300°\text{K}$, with $\varepsilon_0 = 3\text{ eV}$, $\varepsilon_0/k \sim 3 \times 10^4$, this heat capacity contribution is only about one per cent of the value $\frac{3}{2}k$ per electron in a *classical* perfect electron gas.

It was Drude,† at the turn of the century, who first suggested and showed that many of the properties of metals could be explained by assuming the electrons to behave like a free gas in the metal. Others, particularly Lorentz, refined and extended the postulate. The outstanding success of the theory was that it satisfactorily gave the numerical ratio of the heat conductivity of metals to the electrical conductivity, which is (almost) a universal constant proportional to temperature, a fact which is known as the Wiedemann–Franz law. The outstanding failure of the theory was the inability to account for the fact that the free electron gas contributed little or nothing to the heat capacity, since the observed heat capacity of metals was in good numerical agreement with a theory based on the effect of the heavy atoms alone.

Sommerfeld‡ realized that the then recently developed Fermi–Dirac statistics removed this, as well as other difficulties, and with the use of the model of a Fermi–Dirac electron gas he, and others, satisfactorily established the extremely good agreement between the experimental facts and the predictions of theory.

The expression of eq. (2) for the heat capacity, C_{elec}, is linear in temperature at all reasonable temperatures. The lattice contribution, C_{lattice}, from the heavier ions is proportional to T^3 at low temperatures, so that although the absolute electronic contribution decreases with lowered T its relative contribution $C_{\text{elec}}/C_{\text{lattice}}$ goes as T^{-2}. The electronic contribution is measurable at low temperatures and eq. (2) is essentially confirmed experimentally.

5.2. Crystal lattice Hamiltonian

A crystal is essentially a polyatomic molecule of macroscopic size in all three dimensions. In gravity free space it would have 3 translational degrees of freedom, 3 of rotation, and $3N - 6 = 3N(1 - 2N^{-1})$ vibrational degrees of freedom.

---

† P. Drude, *Ann. Physik* **1**, 566 (1900).

‡ A. Sommerfeld, *Z. Physik* **47**, 1 (1928).

Use $x_{\alpha i}$ for the displacement of the atom i, $1 \leqslant i \leqslant N$, from its equilibrium lattice position, along the x_α axis, $\alpha = 1, 2, 3$. The equilibrium positions are defined as those for which the potential energy, U, is a minimum, so that

$$[\partial U/\partial x_{\alpha i}]_{x=0} = 0. \tag{5.2.1}$$

As for the polyatomic molecule we write the potential U as the value, U_0, with all atoms at equilibrium, plus a $\Delta U(x)$ and develop ΔU in powers of the displacements $x_{\alpha i}$, namely with (1),

$$\Delta U(x) = \tfrac{1}{2} \sum_{\alpha i=1}^{3N} \sum_{j\beta=1}^{3N} k_{\alpha i,\beta j} x_{\alpha i} x_{\beta j} + \ldots, \tag{5.2.2}$$

terminating the development with the quadratic terms. The real matrix $\boldsymbol{K}$ of elements $k_{\alpha i,\beta j} = k_{\beta j,\alpha i}$ is symmetric, and for a stable crystal it must be positive definite.[†] The kinetic energy is

$$\text{K.E.} = \sum_{\alpha i=1}^{\alpha i=3N} \tfrac{1}{2} m_i (\dot{x}_{\alpha i})^2. \tag{5.2.3}$$

As in section 3.6 we can find coordinates q_ν for which both (2) and (3) have only diagonal terms and no mixed products. We go through the detailed steps here.

First define normalized coordinates of dimension $m^{1/2}l$ directed along the cartesian axes.

$$y_{\alpha i} = \sqrt{m_i}\, x_{\alpha i}, \tag{5.2.4}$$

so that

$$\text{K.E.} = \tfrac{1}{2} \sum_{\alpha i=1}^{\alpha i=N} (\dot{y}_{\alpha i})^2 \tag{5.2.5}$$

and the momenta conjugate to the $y_{\alpha i}$'s are

$$p_{\alpha i}^{(y)} = \partial(\text{K.E.})/\partial \dot{y}_{\alpha i} = \dot{y}_{\alpha i}. \tag{5.2.6}$$

The different $x_{\alpha i}$'s, and also the different $y_{\alpha i}$'s, are mutually orthogonal, $x_{\alpha i} \cdot x_{\beta j} = y_{\alpha i} \cdot y_{\beta j} = 0$ when $\alpha i \neq \beta j$. Let $\boldsymbol{y}_{\alpha i}^{(0)}$ be unit vectors along these axes,

$$\boldsymbol{y}_{\alpha i}^{(0)} \cdot \boldsymbol{y}_{\beta j}^{(0)} = \delta(\alpha i - \beta j), \tag{5.2.7}$$

where $\delta(\alpha i - \beta j)$ is the Kronecker delta symbol, zero when $\alpha i \neq \beta j$ and

---

[†] Actually, of course, 6 eigenvalues, namely those due to translation and rotation, will be zero, unless terms due to the surroundings are included in U.

unity if $\alpha i = \beta j$. Now define a new set $\boldsymbol{q}^{(0)}$ of $3N$ orthogonal unit vectors in the $3N$-dimensional cartesian space,

$\boldsymbol{q}^{(0)} = \boldsymbol{q}_1^{(0)}, \boldsymbol{q}_2^{(0)}, \ldots, \boldsymbol{q}_\nu^{(0)}, \ldots, \boldsymbol{q}_{3N}^{(0)}$ by $3N$ equations,

$$\boldsymbol{q}_\nu^{(0)} = \sum_{\alpha i=1}^{\alpha i=3N} A_{\nu,\alpha i}\boldsymbol{y}_{\alpha i}^{(0)}, \tag{5.2.8}$$

and require that they be mutually orthogonal and normalized to unit length

$$\boldsymbol{q}_\nu^{(0)} \cdot \boldsymbol{q}_\mu^{(0)} = \delta(\nu - \mu). \tag{5.2.9}$$

One has, from (7) and (8) that

$$A_{\nu,\alpha i} = \boldsymbol{y}_{\alpha i}^{(0)} \cdot \boldsymbol{q}_\nu^{(0)}. \tag{5 2.10}$$

The $y_{\alpha i}^{(0)}$ are linear combinations of the $q_\nu^{(0)}$'s

$$\boldsymbol{y}_{\alpha i}^{(0)} = \sum_{\nu=1}^{\nu=3N} A_{\alpha i,\nu}^{(t)}\boldsymbol{q}_\nu^{(0)}, \tag{5.2.11}$$

and from (9) in (11)

$$A_{\alpha i,\nu}^{(t)} = \boldsymbol{q}_\nu^{(0)} \cdot y_{\alpha i}^{(0)} = A_{\nu,\alpha i}. \tag{5.2.12}$$

Now using (8) for both $\boldsymbol{q}_\nu^{(0)}$ and $\boldsymbol{q}_\mu^{(0)}$ in (9) one has from (7) that

$$\sum_{\alpha i=1}^{\alpha i=3N} A_{\nu,\alpha i}A_{\mu,\alpha i} = \sum_{\alpha i=1}^{\alpha i=3N} A_{\nu,\alpha i}A_{\alpha i,\mu}^{(t)} = \delta(\nu - \mu), \tag{5.2.13}$$

whereas with (11) in (7) using (9),

$$\sum_{\nu=1}^{\nu=3N} A_{\alpha i,\nu}^{(t)}\, A_{\nu\beta j} = \delta(\alpha i - \beta j). \tag{5.2.14}$$

If we define a unit matrix, $\mathbf{A}$, of elements $A_{\nu,\alpha i}$, and its transpose $\mathbf{A}^{(t)}$ of elements $A_{\alpha i,\nu}^{(t)} = A_{\nu,\alpha i}$ the condition of unitarity, which is

$$\mathbf{A} \cdot \mathbf{A}^{(t)} = \mathbf{A}^{(t)} \cdot \mathbf{A} = \mathbf{1}, \tag{5.2.15}$$

with $\mathbf{1}$ the unit matrix, is just the matrix notation for the two sets of equations, (13) and (14). We define

$$K_{\nu,\mu} = \sum_{\alpha i=1}^{\alpha i=3N} \sum_{\beta j=1}^{\beta j=3N} A_{\nu,\alpha i}\left[\frac{k_{\alpha i,\beta j}}{m_i^{1/2}\, m_j^{1/2}}\right] A_{\beta j,\mu}^{(t)}. \tag{5.2.16}$$

Let $\boldsymbol{q}_\nu, \boldsymbol{q}_\mu$ be the displacements from equilibrium along the unit vectors $\boldsymbol{q}_\nu^{(0)}$, $\boldsymbol{q}_\mu^{(0)}$. With (4) and (8) to replace the displacements $x_{\alpha i}$, $x_{\beta j}$ in (2) by $\boldsymbol{q}_\nu$, $\boldsymbol{q}_\mu$, we have

$$\Delta U(\boldsymbol{q}) = \tfrac{1}{2} \sum_{\mu=1}^{\mu=3N} \sum_{\nu=1}^{\nu=3N} q_\nu K_{\nu\mu} q_\mu. \tag{5.2.17}$$

Now there is a fundamental theorem that the symmetric matrix $\boldsymbol{K}^{(y)}$ of elements

$$K_{\alpha i,\beta j}{}^{(y)} = \frac{k_{\alpha i,\beta i}}{m_i{}^{1/2}m_j{}^{1/2}} = K_{\beta j,\alpha i}{}^{(y)} \tag{5.2.18}$$

can be brought into a "Diagonal Representation" by a unitary transform, namely that we can choose the $A_{\nu,\alpha i}$, consistent with (12), (13) and (14) so that $K_{\nu,\mu}$ of (16) is zero if $\nu \neq \mu$.

To accomplish this multiply eq. (16) on both sides by $A_{\mu,\beta j}$ and sum over ν. If, now

$$K_{\nu,\mu} = K_{\mu,\mu}\,\delta(\nu - \mu), \tag{5.2.19}$$

use (14) to find,

$$K_{\mu\mu}A_{\mu,\beta j} = \sum_{\alpha i=1}^{\alpha i=3N} A_{\mu\alpha i}K_{\alpha i,\beta j}. \tag{5.2.20}$$

These form, for each μ value, $3N$ equations with $1 \leqslant \beta j \leqslant 3N$ for the $3N$ unknowns $A_{\mu,\alpha i}$,

$$\sum_{\alpha i=1}^{\alpha i=3N} [K_{\alpha i,\beta i} - \sigma(\alpha i - \beta j)K_{\mu\mu}]A_{\mu,\beta j} = 0. \tag{5.2.21}$$

These equations have non-trivial solutions, $A_{\mu,\beta j} \neq 0$ *only* if the determinant of the coefficients is zero,

$$\begin{vmatrix} (K_{11,11} - \gamma)\ K_{21,11} \cdot \cdot \cdot K_{3N,11} \\ K_{11,21}\ (K_{21,21} - \gamma) \cdot \cdot \cdot K_{3N,21} \\ \cdot \\ \cdot \\ \cdot \\ K_{11,3N} \cdot \cdot \cdot \qquad (K_{3N,3N} - \gamma) \end{vmatrix} = 0. \tag{5.2.22}$$

The determinantal equation of (22) is that a polynomial in γ of order $3N$ be zero. This has, in general $3N$ roots, the $3N$ values of $K_{\mu\mu}$. Six of the roots, corresponding to coordinates of translation and rotation will be zero. For a stable crystal the other $3N - 6$ will be positive. With (5) for the kinetic energy in terms of the $\dot{y}_{\alpha i}$'s, using, from (11) that

$$\text{K.E.} = \tfrac{1}{2} \sum_{\nu=1}^{\nu=3N} \sum_{\mu=1}^{\mu=3N} [\sum_{\alpha i=1}^{\alpha i=3N} A_{\nu,\alpha i}A_{\alpha i,\mu}{}^{(t)}]\dot{q}_\nu\dot{q}_\mu, \tag{5.2.23}$$

which, from (13), is,

$$\text{K.E.} = \tfrac{1}{2} \sum_{\mu=1}^{\mu=3N} (\dot{q}_\mu)^2, \qquad (5.2.24)$$

so that, by its definition,

$$p_\mu = \partial(\text{K.E.})/\partial\dot{q}_\mu = \dot{q}_\mu. \qquad (5.2.25)$$

Finally with (25) in (24) for the kinetic energy and with (19) in (17) for the potential energy we have that the Hamiltonian is separable,

$$H = \sum_{\mu \geqslant 1}^{\mu=3N} H_\mu(p_\mu, q_\mu), \qquad (5.2.26)$$

$$H_\mu(p_\mu, q_\mu) = \tfrac{1}{2}[p_\mu{}^2 + K_{\mu\mu}\, q_\mu{}^2]. \qquad (5.2.27)$$

The classical equations of motion are

$$\dot{q}_\mu = \partial H/\partial p_\mu = p_\mu, \qquad (5.2.28)$$

which repeats (25) and

$$\dot{p}_\mu = \ddot{q}_\mu = -\partial H/\partial q_\mu = -K_{\mu\mu} q_\mu, \qquad (5.2.29)$$

which has the solution

$$q_\mu = a_\mu \cos(2\pi\nu_\mu t + \sigma), \qquad (5.2.30)$$

with the frequency ν_μ given by

$$\nu_\mu = (1/2\pi)\sqrt{K_{\mu\mu}}. \qquad (5.2.31)$$

The motion of the system can be analysed as independent harmonic vibrations along the "normal coordinates" q_μ with frequency ν_μ.

5.3. Lattice vibration spectra

As is discussed in section 3.5 the average energy of a single classical harmonic oscillator degree of freedom is kT, and the heat capacity is k. Since the lattice motions can be analysed as $3N$ independent harmonic oscillator degrees of freedom one predicts for the heat capacity, C_V, at constant volume, that in the classical regime,

$$C_V = (\partial E/\partial T)_V = 3N\, kT, \qquad (5.3.1)$$

for a crystal having N atoms. The expression of eq. (1) is that of the law of Dulong and Petit. It holds quite precisely for the heavy metal elements, and also for many simple inorganic compounds, at room temperature.

However, it was known quite early that some light elements, particularly diamond, had a far lower heat capacity than eq. (1) predicts, and that quite generally all heat capacities tended toward zero as the temperature tended toward 0°K. Einstein† was the first to recognize that the model used by Planck to explain Black Body Radiation (section 7.3) required quantized oscillator levels in any material oscillator. He assumed a single frequency for each of the $3N$ degrees of freedom in the crystal. The energy of N_0 oscillators of frequency ν is given in section 3.10g, and for $3N$ it is

$$E_{\text{Einstein}} = 3NkT \frac{h\nu/kT}{e^{h\nu/kT} - 1}, \tag{5.3.2}$$

$$C_{V_{\text{Einstein}}} = 3Nk\left(\frac{h\nu}{kT}\right)^2 [e^{h\nu/kT} - 1]^{-1}[1 - e^{-h\nu/kT}]^{-1}. \tag{5.3.3}$$

These are known as the Einstein equations for energy and heat capacity respectively.

The expression of (3) did indeed predict a much lower C_V-value than eq. (1) for $h\nu/kT > 1$, but predicted an exponential decrease of C_V as $T \rightarrow 0$. The actual decrease was known to be much slower, namely as T^3.

Born and Karman‡ actually computed the spectra of $K_{\mu\mu}$ values for a *one-dimensional*, one component crystal with nearest neighbor interaction. They also deducted that for three dimensions the number of eigenvalues $K_{\mu\mu}$ leading to frequencies between ν and $\nu + d\nu$ would, at low frequencies, increase as $\nu^2 d\nu$. However, at the same time Debye§ gave his remarkably simple and ingenious analysis that subsequently dominated thinking about the spectra of frequencies in crystals. The Debye equation was hardly improved upon for about a quarter century. The Debye equations will be discussed in the next section.

Many recent computer calculations and many recent experimental techniques have led to a tremendously increased knowledge|| of the lattice vibrational spectra for a wide variety of crystals. The concepts of Debye are, however, so simple and so enlightening that we discuss them in more detail.

---

† Albert Einstein, *Ann. Physik* **22**, [4], 180 (1907).

‡ M. Born and Th. von Karman, *Physik Z.* **13**, 297 (1912).

§ P. P. Debye, *Ann. Physik* **39**, 789 (1912).

|| See for instance *Lattice Dynamics*, edited by R. F. Wallis, Pergamon Press, 1965.

5.4. The Debye lattice theory

An elastic continuum can support standing oscillatory waves of wavelength λ, and the amplitudes of these standing waves correspond, in the continuous medium, to the displacements q_ν along the normal coordinate vectors $\boldsymbol{q}_\nu^{(0)}$ discussed in section 2. The Debye idea is to use these normal coordinates and terminate them at a maximum frequency ν_m such that the total number of degrees of freedom has the required value of $3N$. In addition Debye postulates a constant sound velocity.

The displacement vector, $\boldsymbol{\psi}(\boldsymbol{r})$, of the matter at position $\boldsymbol{r}$ from its equilibrium position can be given as a sum of Fourier components. With λ_x the wavelength along the x-axis, and with

$$k_x = \lambda_x^{-1}, \qquad k_y = \lambda_y^{-1}, \qquad k_z = \lambda_z^{-1}, \tag{5.4.1}$$

$$\kappa = [k_x^2 + k_y^2 + k_z^2], \tag{5.4.2}$$

there are

$$N_\kappa(\kappa)d\kappa = 4\pi V\kappa^2 d\kappa, \tag{5.4.3}$$

orthogonal wave functions of wave number κ between κ and $\kappa + \mathrm{d}\kappa$. There are, then, this many degrees of freedom in compressional waves between κ and $\kappa + \mathrm{d}\kappa$ and twice that number in transverse waves, since there are two mutually perpendicular directions of the displacement vector $\boldsymbol{\psi}(\boldsymbol{r})$, both perpendicular to the wave vector $\boldsymbol{k}$ of components k_x, k_y, k_z.

In general the wave frequency ν is $\nu = \kappa c$ with c the wave velocity.

The Debye approximation is obtained by making two assumptions which are certainly not exact. The first of these is to assume that the wave velocity c for both longitudinal and transverse waves is independent of wave number. We then have

$$\kappa^2 d\kappa = c^{-3}\nu^2 d\nu, \tag{5.4.4}$$

so that with c_l the sound velocity, and c_t the velocity of the transverse wave, the number $N_\nu(\nu)d\nu$ of degrees of freedom between ν and $\nu + d\nu$, will be

$$N(\nu)d\nu = 4\pi V\left[\frac{2}{c_t^3} + \frac{1}{c_l^3}\right]\nu^2 d\nu. \tag{5.4.5}$$

The second assumption is that (5) holds up to a single maximum frequency ν_m. The total number of degrees of freedom is $3N$. Eliminate the wave velocities from (5) by writing $N(\nu) = AN\nu^2 d\nu$ and determine A by the value, ν_m, of the maximum frequency,

$$\int_0^{\nu_m} N(\nu)d\nu = AN\int_0^{\nu_m} \nu^2 d\nu = \frac{1}{3} AN{\nu_m}^3 = 3N, \qquad (5.4.6)$$

$$A = 9/{\nu_m}^3, \qquad (5.4.7)$$

so that
$$N(\nu)d\nu = 9N(\nu/\nu_m)^2\, d(\nu/\nu_m). \qquad (5.4.8)$$

Use $kT(h\nu/kT)[\exp(h\nu/kT) - 1]^{-1}$ for the energy at T of each degree of freedom above the lowest state. Define

$$\theta = h\nu_m/k, \qquad (5.4.9)$$

as the "Debye theta," the energy of the maximum frequency, ν_m in units of k, and with

$$D(\theta/T) = 3\frac{T^3}{\theta}\int_0^{\theta/T} \frac{x^3 dx}{e^x - 1} \qquad (5.4.10)$$

we have for the crystal energy E, above that at $T = 0$,

$$E(T) - E_0 = 3NkT\, D(\theta/T). \qquad (5.4.11)$$

The Debye function, $D(u)$, may be evaluated approximately in different manners for high temperatures, $u = \theta/T$ small, and for low temperatures, $u = \theta/T$ large. For small u, with $x < u$ we develop

$$e^x = 1 + x + \tfrac{1}{2}x^2 + \ldots \text{ so that}$$

$$x^3[e^x - 1]^{-1} = x^3\left[x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \ldots\right]^{-1}$$

$$= x^2 - \frac{1}{2}x^3 + \frac{1}{12}x^4 - \frac{1}{720}x^6,$$

and on integration one finds from (10) that

$$D(u) = 1 - \frac{3}{8}u + \frac{1}{20}u^2 - \frac{1}{1680}u^4. \qquad (5.4.12)$$

To evaluate $D(u)$ for large u-values we first write

$$\frac{x^3}{e^x - 1} = \frac{e^{-x}x^3}{1 - e^x} = \sum_{n\geqslant 1} x^3 e^{-nx}\, dx$$

and then integrate by

$$\int_0^u \frac{x^3}{e^x - 1}\,dx = \sum_{n\geqslant 1}\left\{\int_0^\infty x^3 e^{-nx}\,dx - \int_u^\infty x^3 e^{-nx}\,dx\right\}$$

$$= 6\sum_{n\geqslant 1}^{\infty} n^{-4} - (u^3 + 3u^2 + 6u + 6)e^{-\mu}$$

$$+ (\tfrac{1}{2}u^3 + \tfrac{3}{4}u^2 + \tfrac{3}{4}u + \tfrac{3}{8})\, e^{-2\mu} + \ldots.$$

With $\Sigma n^{-4} = \pi^4/90$ one finds from (10) that

$$D(u) = \frac{\pi^4}{5}\,u^{-3} - (3 + 9u^{-1} + 18u^{-2} + 18u^{-3})e^{-u}$$

$$- \tfrac{3}{8}(4 + 6u^{-1} + 6u^{-2} + 3u^{-3})e^{-2-} + \ldots. \qquad (5.4.13)$$

Using (12) and (13) in eq. (11) one obtains

$$E(T) - E_0 = 3NkT\left[1 - \frac{3}{8}\frac{\theta}{T} + \frac{1}{20}\left(\frac{\theta}{T}\right)^2 - \frac{1}{1680}\left(\frac{\theta}{T}\right)^4 + \ldots\right],$$

$$\frac{\theta}{T} < 1, \qquad (5.4.14)$$

$$E(T) - E_0 = \frac{3}{5}\,\pi^4(Nk\theta)(T/\theta)^4,\ (\theta/T >> 1). \qquad (5.4.15)$$

The energy of the crystal measured from E_0, the energy at $T = 0$, approaches the classical linear slope of $3Nk$ against T, but remains below the classical value of $3NkT$ by the constant term $(9/8)(Nk\theta)$.

If, however, the energy is measured from the minimum of the potential energy of the crystal then the energy $\frac{1}{2}h\nu$ per degree of freedom must be added to (11). This amount is, from (8),

$$\int^{\nu_m} \frac{1}{2}\,h\nu N(\nu)d\nu = \frac{9}{8}\,N\,h\nu_m = \frac{9}{8}\,(Nk\theta). \qquad (5.4.16)$$

Add this to (14) to find

$$E - U_0 = 3NkT\left[1 + \frac{1}{20}\left(\frac{\theta}{T}\right)^2 - \frac{1}{1680}\left(\frac{\theta}{T}\right)^4 + \ldots\right], \qquad (5.4.17)$$

approaching the classical value from above as T increases. The quantum mechanical energy at absolute zero, E_0, is higher than the classical value by $(9/8)(Nk\theta)$. As the temperature increases the energy increases more slowly than the classical linear $3NkT$ and asymptotically approaches the classical value as T^{-2}.

At low temperatures the Debye energy goes as T^4,

$$E(T) - E_0 = \frac{3}{5}\pi^4(Nk\theta)(T/\theta)^4, \tag{5.4.18}$$

and the low temperature heat capacity as

$$C_V = (\partial E/\partial T)_V = \frac{12}{5}\pi^4 Nk(T/0)^3. \tag{5.4.19}$$

Experimental heat capacities of regular monatomic crystals fit the Debye curve quite well, if the experimentally measured C_p is corrected to C_V. However, the fit is by no means perfect. The low temperature region proportional to T^3 if fitted to (19) carefully does not lead to the same value of θ that makes the best fit of the intermediate temperature range.

At sufficiently low frequencies, long wavelengths, the crystal would be expected to behave like a continuum, and the restoring force of any displacement from equilibrium should be proportional only to the second derivative of the displacement amplitude. This is the essential requirement for the wave velocity to be constant. For an isotropic crystal the wave velocities should also not depend on the direction of propagation. As the wavelengths approach a small integer times the interatomic distances one cannot expect that the wave velocities be constant, nor independent of the direction of the wave vector. Neither is the limiting wave vector magnitude at which the sine function should be terminated independent of the direction of the vector k, nor will it occur at the same frequency for the waves of different velocities. The Debye heat capacity curve is a useful approximation, but not a function which should be expected to fit with high precision.

5.5. Debye thermodynamic functions

The energy of the Debye crystal of N atoms is given by eq. (5.4.11) as

$$E(T) - E_0 = 3NkT\, D(\theta/T). \tag{5.5.1}$$

Since from (5.4.10),

$$T(d/dT)D(\theta/T) = 3\left\{D\left(\frac{\theta}{T}\right) - \frac{\theta}{T}[e^{\theta/T} - 1]^{-1}\right\}, \tag{5.5.2}$$

we have

$$C_V = \partial E/\partial T = 3Nk\left\{4D\left(\frac{\theta}{T}\right) - 3\left(\frac{\theta}{T}\right)[\theta^{\theta/T} - 1]^{-1}\right\}. \tag{5.5.3}$$

At high temperatures this becomes

$$C_V = 3Nk\left[1 - \frac{1}{20}\left(\frac{\theta}{T}\right)^2 + \frac{1}{560}\left(\frac{\theta}{T}\right)^4 - \ldots\right], \tag{5.5.4}$$

which is obtained either by expanding (3), or differentiating (5.4.14). At low temperatures the expression (5.4.16) holds.

The entropy of a single oscillator of frequency ν is, from section 3.10g,

$$S = k\left\{\frac{h\nu}{kT}\,(e^{h\nu/kT} - 1)^{-1} - \ln\,(1 - e^{-h\nu/kT})\right\}. \tag{5.5.5}$$

We multiply this by (5.4.8) and integrate. One finds

$$S = 3Nk\left[D\left(\frac{\theta}{T}\right) - 3\left(\frac{T}{\theta}\right)^3\int_0^{(\theta/T)} x^2\ln\,(1 - e^{-x})dx\right].$$

By partial integration this becomes

$$S = 3Nk\left[\frac{4}{3}D\left(\frac{\theta}{T}\right) - \ln\,(1 - e^{-\theta/T})\right]. \tag{5.5.6}$$

The Helmholtz free energy is $A = E - TS$,

$$A - E_0 = 3NkT\left\{\ln\,(1 - e^{-\theta/T}) - \frac{1}{3}D\left(\frac{\theta}{T}\right)\right\} \tag{5.5.7}$$

if measured from the energy E_0 of the crystal at absolute zero or

$$A - U_0 = 3NkT\left\{\ln\,(1 - e^{-\theta/T}) + \frac{3}{8}\left(\frac{\theta}{T}\right) - \frac{1}{3}D\left(\frac{\theta}{T}\right)\right\} \tag{5.5.8}$$

if measured from the minimum potential.

5.6. Thermal expansion

The expression (5.5.8) for the Helmholtz free energy A depends on the potential energy, U_0, when the atoms are in their equilibrium positions, and on the maximum frequency ν_m through the Debye theta. If the crystal is stressed by tension or compression at the surface the equilibrium distance r_0 between nearest neighbors will change due to the volume change. The potential energy, U_0, of the lattice with the atoms at the lattice sites depends on r_0, and in general, so does the maximum Debye frequency, ν_m. If, then, we write both $U_0(V)$ and $\theta(V)$ we have an expression,

$$A(N,V,T) = U_0(V) + 3NkT\left\{\ln\,(1 - e^{-\theta/T}) + \frac{3}{8}\left(\frac{\theta}{T}\right) - \frac{1}{3}D\left(\frac{\theta}{T}\right)\right\}, \tag{5.6.1}$$

of A as a function of both volume and temperature.

The pressure, P, is $-(\partial A/\partial V)_T$, so that in principle eq. (1) determines the pressure and therefore the equation of state. Since the relations that we wish to derive are more general than the Debye theory we will rewrite A in a more general form. The Helmholtz free energy, a, per oscillator degree of freedom is $a = kT \ln\,(1 - e^{-h\nu/kT})$. Assume that the number of degrees of freedom of the crystal of N atoms between ν and $\nu + d\nu$, is of the form

$$N(\nu)d\nu = 3Nf(\nu/\nu_0)d(\nu/\nu_0), \tag{5.6.2}$$

where ν_0 is some characteristic frequency dependent on the volume only. With $\theta = h\nu_0/k = \theta(V)$, and $x = \nu/\nu_0$, one has

$$\begin{aligned} A(N,V,T) &= U_0(V) + 3NkT \int f(x) \ln\,(1 - e^{-\theta x/T})dx \\ &= U_0(V) + 3NkT\, G(\theta/T), \end{aligned} \tag{5.6.3}$$

in agreement with (1) in which the function $G(\theta/T)$ is given explicitly for a Debye crystal.

Now write,

$$\begin{aligned} A - T(\partial A/\partial T)_V &= A + TS = E \\ &= -\,3NkT[T\,\partial G(\theta/T)/\partial T]_V, \end{aligned} \tag{5.6.4}$$

which can be explicitly checked for eq. (1) with (11) and (13). On the other hand

$$\begin{aligned} PV &= -V(\partial A/\partial V)_T = \\ &-V\,dU_0/dV - 3NkT[V\,\partial G(\theta/T)/\partial V]_T. \end{aligned} \tag{5.6.5}$$

The function G depends on V through $\theta = \theta(V)$ only, and since G depends on θ/T only

$$\begin{aligned} [V(\partial G/\partial V]_T &= -(d \ln \theta/\ln V)(T\,\partial G/\partial T)_V \\ &(d \ln \theta/d \ln V)(E/3NkT), \end{aligned} \tag{5.6.6}$$

so that

$$P = -\,(dU_0/dV) - E(d \ln \theta/dV). \tag{5.6.7}$$

Differentiate eq. (7) with respect to temperature at constant volume. With $\kappa = -V^{-1}(\partial V/\partial P)_T$ the isothermal compressibility and $\alpha = V^{-1}(\partial V/\partial T)_P$ the thermal expansion coefficient we have

$$\left(\frac{\partial P}{\partial T}\right)_V = \frac{\alpha}{\kappa} = -\left(\frac{\partial E}{\partial T}\right)_V V^{-1}\left(\frac{d \ln \theta}{d \ln V}\right),$$

$$\alpha = -(\kappa C_V/V)(d \ln \theta/d \ln V). \qquad (5.6.8)$$

The thermal expansion coefficient should be proportional to the heat capacity, a prediction which is verified experimentally.

5.7. Crystal imperfections and order–disorder

It is relatively easy to show that at equilibrium there must always be a non-zero concentration of defects in crystals. The defects, in general, are places in which the energy is higher than that of the perfect lattice by some amount ε, and their concentration will be of order $\exp - \varepsilon/kT$.

A simple derivation of this may be instructive. Consider N lattice sites, almost all of which are occupied by atoms in a "normal" manner, but with N_0 abnormal sites of excess energy ε. The N_0 abnormalities can be distributed on the N sites in $N!/[(N - N_0)!N_0!]$ different ways, adding an excess entropy,

$$\Delta S = k[\ln N! - \ln (N - N_0)! - \ln N_0!], \qquad (5.7.1)$$

$$\begin{aligned} d\Delta S/dN_0 &= k \ln[(N - N_0)/N_0] \\ &\simeq k \ln (N/N_0), \end{aligned} \qquad (5.7.2)$$

when $N_0 << N$. The value of $d\Delta S/dN_0$ becomes logarithmically infinite as $N_0/N \to 0$. The value of the increase in energy,

$$dE/dN_0 = \varepsilon, \qquad (5.7.3)$$

per increase in the number, N_0, of abnormalities is finite, and, as long as the defect concentration remains low enough that they do not interact with each other, it is independent of N_0/N. The Helmholtz free energy is minimized when

$$(\partial A/\partial N_0)_T = 0 = [\partial(E - TS)/\partial N_0]_T = \varepsilon - kT \ln (N/N_0),$$

or, at equilibrium,

$$N_0/N = \exp - (\varepsilon/kT). \qquad (5.7.4)$$

Actually in many crystals the energy ε necessary to produce a defect is so large compared with kT that N_0/N is truly negligible. For instance the energy necessary to produce a vacancy in the lattice is of that order necessary to vaporize an atom. The concentration of vacancies at equilibrium will be approximately that of the atomic gas in equilibrium with the crystal, and this may well be much less than one atom per cm^3.

However, in a considerable number of crystals there is an inherent disorder above some characteristic temperature. Typical cases of this are solid solutions in which one type of lattice site can be occupied at random by either of two species of atom or ion. Other cases include the random up-down arrangement of the CO-molecule in its crystal, a randomness that arises from the very small asymmetry in the molecule itself. A similar rotational randomness is found in the crystals of many organic molecules that have nearly spherical or cylindrical symmetry. This latter type of randomness has been described as "free rotation," a term which is probably unfortunate even when the transitions between different orientations may be extremely fast if measured in units of seconds.

Presumably in all cases the randomness would disappear at equilibrium at $T \to 0$, and the entropy at equilibrium and at zero temperature would vanish. In some cases, however, the randomness becomes "frozen in." In these cases the rate of relaxation to the true equilibrium of the lowest energy configuration is so slow below the temperature at which the ordered arrangement becomes stable that, in practice, the disordered form persists to zero temperature, and the observed entropy extrapolated to $T = 0$ is non-zero. In other cases equilibrium is established reasonably rapidly. An ordered form exists at low temperature, and a disordered form at high temperatures.

The transition between the two cases may be, even with true temperature equilibrium at all stages, gradual on a temperature scale, or discontinuous; the "gradual" cases showing varying degrees of abruptness.

The extreme of the discontinuous case is a first order change of crystal structure from a crystal of one symmetry at the low temperature to one of higher symmetry at the higher temperature. The higher symmetry crystalline form then permits lattice site occupancy in random patterns. There is little to say quantitatively about such a case. The energy of the high symmetry, high temperature, form is higher than that of the low temperature form, but the randomness introduces an additional entropy term. As T increases the greater entropy increases in importance in the free energy, $A = E - TS$, and the higher energy, higher entropy crystal becomes the more stable.

The case of rather abrupt transition which is manifested by a noticeable peak in the heat capacity curve, goes under the general name of a "cooperative phenomenon." Actually the most extreme cooperative case is that of the first order transition, a complete change of phase, but the term is usually reserved for the less extreme but still marked continuous transition.

A caricature of a cooperative phenomenon can be described as follows. Suppose in a system there are n units that can exist in two energy levels, the first of zero energy having degeneracy g_0 and the second of energy ε having degeneracy g_1. The energy contribution from these to the energy of the system is

$$\begin{aligned}\Delta E(T) &= ng_1\, \varepsilon\, e^{-\varepsilon/kT}[g_0 + g_1\, e^{-\varepsilon/kT}]^{-1}\\ &= n\varepsilon[1 + (g_0/g_1)e^{\varepsilon/kT}]^{-1},\end{aligned} \tag{5.7.5}$$

and the contribution to the heat capacity is

$$C(T) = nkg_1g_0(\varepsilon/kT)^2[g_0e^{\varepsilon/kT} + g_1]^{-2}\, e^{\varepsilon/kT}. \tag{5.7.6}$$

This has a maximum at

$$\begin{aligned}&Td\Delta C/dT = 0\\ &= nk\, g_1g_0(\varepsilon/kT)^2\, e^{\varepsilon/kT}[g_0e^{\varepsilon/kT} + g_1]^{-3}\left[g_0e^{\varepsilon/kT}\left(\frac{\varepsilon}{kT} - 2\right) -\right.\\ &\left. g_1\left(\frac{\varepsilon}{kT} + 2\right)\right], \qquad \exp\frac{\varepsilon}{kT_m} = \frac{g_1}{g_0}\frac{[1 + (2kT_m/\varepsilon)]}{[1 - (2kT_m/\varepsilon)]}.\end{aligned} \tag{5.7.7}$$

Replace $e^{\varepsilon/kT}$ by this in ΔC for T_m and we find

$$\Delta C_m = nk\left(\frac{\varepsilon}{2kT_m}\right)^2\left[1 + \frac{2kT_m}{\varepsilon}\right]\left[1 - \frac{pkT_m}{\varepsilon}\right]. \tag{5.7.8}$$

If $g_0 = g_1 = 1$ the value of T_m turns out to be about $0.42(\varepsilon/k)$ and ΔC_m about $0.42k$. At $T = \frac{1}{2}T_m$ and at $T = 2T_m \Delta C_m$ still has about half its maximum value. If, however, $g_1 >> g_0$ then T_m is much lower, or, conversely if T_m occurs at the same temperature the energy ε is much larger. The extreme which leads to a first order transition is that $(g_1/g_0) \sim \exp N$ with N the number of molecules in the system, $N \sim 10^{23}$. In this case $\varepsilon \sim NkT_m$, the maximum of $\Delta C/nk$ is of the order of N^2 and with half-width of order N^{-1}. If, however, $g_1/g_0 = 10^4$ one finds $\varepsilon/kT_m = 9.2$, and $\Delta C_m/nk \sim 20$, while the half-width is in a range of $\Delta T \sim 10^{-1}T_m$.

Independent atoms, or independent defects which can occur at any lattice site correspond to these equations with $n = N$, the total number of sites in the crystal lattice, and g_0 and g_1 of the order of unity or a small integer. In the first order transition the unit that transforms is the whole system. We can use the equations (5) to (8) with $n = 1$, but g_1 is of order $\exp N$. The typical "cooperative phenomenon" is one in which the behavior at the neighboring or near neighboring lattice sites are correlated. When one site is excited it becomes easier for those around it to become excited. The unit that becomes excited is much larger than a single atom. The degeneracy, g_1, of the exciton, which, with ΔS the entropy increase in this excited unit, is $g_1 = \exp \Delta S/k$, can be of order several powers of ten, but is not of order $\exp N$. The heat capacity curve may be extremely sharply peaked, but is not necessarily singular as the peak is passed.

There are, however, truly singular heat capacity curves which do not correspond to first order phase transitions, but to higher orders, namely to a discontinuity in the nth derivative with respect to T or P of the Gibbs free energy. A discontinuity in the first derivatives, $(\partial G/\partial T)_p = -S$ or $(\partial G/\partial P)_T = V$ is a first order transition. One in the second derivatives $(\partial^2 G/\partial T^2)_P = -(\partial S/\partial T)_P = -C_P/T$, $(\partial^2 G/\partial P^2)_T = (\partial V/\partial P)_T = -V\kappa$, means a discontinuity in the heat capacity and compressibility and was called a second order phase transition by Ehrenfest. Presumably third order transitions could exist but they would be very difficult to establish with certainty. Even second order transitions are experimentally difficult to verify.

Finally, one interesting observation may be made about the crystals that show transitions due to disordered rotational configurations of the molecules. Simple monatomic crystals have an entropy of melting of approximately Boltzmann's constant k per atom. Crystals of structured molecules tend to have appreciably higher entropies of melting, since the liquid allows not only greater freedom of position of the center of mass coordinates, but also of the angular coordinates of orientation. If the rotational transitions are relatively sharp their entropy increase can be approximately evaluated by integrating $\Delta C/T$ through the transition. It is generally found that whereas the entropies of melting crystals have the value characteristic of monatomic crystals, the sum of the entropies of the crystal transitions with that of melting corresponds, approximately, to the entropy change expected for melting of crystals of molecules with similar structure.

CHAPTER 6

DENSE FLUIDS

6.1. Notation and the potential energy

For the perfect gas the fact that the Hamiltonian is separable in the coordinates of the individual molecules permits easy evaluation of the partition functions. The treatment by virial developments, of the imperfections of a mildly imperfect gas is essentially a perturbation treatment from this starting point. At the other extreme of high densities and low temperatures only configurations close to the crystalline configuration of minimum energy are important. The approximation of using only the quadratic terms of the development of the potential energy in terms of displacements again leads to a separable Hamiltonian, and again evaluation of the partition function is relatively easy.

The intermediate range of moderately high temperatures and high densities offers much more difficulty. In this range the materials are fluid either as true liquid, or as the highly compressed gas above the critical temperature and also above the critical density, which is usually classed with the liquid phase. In this range the absence of peculiar possibilities of simplification necessitates the use of the most general methods which can be devised. As a result it is not surprising that the methods used are such that they are actually valid in all ranges crystalline, or dilute gaseous, as well as in the liquid. They may not be, and usually are not, the methods of choice in these cases. The methods discussed in this chapter are then completely general, and applicable to all thermodynamic phases.

We will, however, at the outset make one simplification, which, in practice at least, usually limits the applicability to liquids and gases, and excludes crystals at very low temperatures. This simplification is to assume that the classical limit applies. Actually even this simplification is unnecessary, and only one section, section 6.4 in which we discuss the potentials of average force, would be invalid. If the diagonal part of the coordinate representation of the quantum mechanical density matrix is used to replace $\exp - U/kT$, then all the other manipulations of this chapter remain valid. With this one dogmatic remark about quantum statistics we shall proceed to forget that not all systems are classical.

The classical distribution in the cartesian momentum space is always Gaussian, proportional to $\exp -(\boldsymbol{p}\cdot\boldsymbol{p}/2mkT)$, independent of the coordinates. Our whole concern, then, is with the coordinate configuration. The term $[\exp -(PV/kT)]\,(z^N/N!)\exp -(U/kT)$ gives the probability density that there are exactly the set, N, of molecules in the volume V, and that these are at the configurational positions for which U is the potential energy. This is the most detailed information that could conceivably be given about the structure in configuration space of an ensemble. Indeed it is too detailed to assimilate. The question arises whether one can ask a more limited question about the structure of the systems of the ensemble; one which goes beyond the limited information given by the thermodynamic functions, but which is still interpretable. The answer lies in the reduced probability density function, ρ_n, which are functions of a limited set of coordinates of n molecules, with $n = 1, 2, 3, \ldots$, etc.

We will first introduce a much needed shorthand notation for the coordinates. The coordinates of the ith molecule may be designated by $\boldsymbol{q}_i$. Although we shall generally describe the systems as though $\boldsymbol{q}_i$ included only the three-vector, $\boldsymbol{r}_i$, of the position of the center of mass of molecule i, the formalism is equally valid if $\boldsymbol{q}_i$ includes internal coordinates as well. In general the nature of these internal coordinates will depend on the species, a, of molecule i. We have previously used $\boldsymbol{N}$ to designate a set, N_a of species a, N_b of species $b, \ldots$ of molecules. We shall also use $\boldsymbol{n}, \boldsymbol{\nu}, \ldots$, etc., to indicate number sets. We will use upper case $\boldsymbol{N}, \boldsymbol{M}$, etc., when the emphasis is on numbers $N \sim 10^{23}$, and lower case $\boldsymbol{n}$, when the emphasis is on $n = 1, 2, 3, \ldots$, etc., although actually the functions introduced will be defined for all numbers, unity to infinity, in all cases. We now use $\{\boldsymbol{N}\}$, (or $\{\boldsymbol{n}\}$ etc.) to indicate the coordinates of this number set,

$$\{\boldsymbol{N}\} = \boldsymbol{q}_{1a}, \ldots, \boldsymbol{q}_{Na}, \boldsymbol{q}_{1b}, \ldots, \boldsymbol{q}_{Nb}, \ldots, \boldsymbol{q}_{Nb}, \ldots \qquad (6.1.1)$$

and for the volume element we use

$$d\{\boldsymbol{N}\} = J\{\boldsymbol{N}\}d\boldsymbol{q}_{1a} \ldots d\boldsymbol{q}_{Na}\, d\boldsymbol{q}_{1b} \ldots. \qquad (6.1.2)$$

The symbol $J\{\boldsymbol{N}\}$ is the Jacobian of the volume element from the cartesian coordinates of the atomic constituents of the molecules to the internal coordinates used. If the $\boldsymbol{q}_i$'s are solely the cartesian $\boldsymbol{r}_i$'s then $J\{\boldsymbol{N}\} \equiv 1$.

We also wish to introduce a notation for the coordinates of a set $\boldsymbol{n}$ of molecules which is a subset of a larger set N. The symbol $[\boldsymbol{n}]_N$ is used for this. For given numbers, $\boldsymbol{n} = n_a, n_b, \ldots$, there are

$$N!/[(N - \boldsymbol{n})!\,\boldsymbol{n}!] \text{ different subsets, } [\boldsymbol{n}]_N \text{ of } [N]. \qquad (6.1.3)$$

In general we include as a possible subset, in our notation, the improper subset $\{N\} = \{\boldsymbol{n}\}_N$ which is the subset of all, so that when summation over all subsets is written the last term is that of the complete set. One may thus write, for the total potential energy, U_N, of the set N,

$$\begin{aligned} U_N\{N\} &= \sum_{\{n\}_N} u_n\{\boldsymbol{n}\}_N \\ &= \sum_a \sum_{ia=1}^{ia=N_a} u_a(\boldsymbol{q}_{ia}) + \sum_a \sum_{N_a = ia > ja \geqslant 1} \sum u_{aa}(\boldsymbol{q}_{ia}, \boldsymbol{q}_{ja}) \\ &\quad + \sum_{a>b}\sum \sum_{ia=1}^{ia=N_a} \sum_{jb=1}^{jb=N_b} u_{ab}(\boldsymbol{q}_{ia}, \boldsymbol{q}_{jb}) + \ldots \end{aligned} \qquad (6.1.4)$$

If the $\boldsymbol{q}_{ia}$ include no internal coordinates, $q_{ia} \equiv r_{ia}$, and in the absence of any external field, we would choose the zero of energy so that $u_a(\boldsymbol{q}_{ia}) \equiv 0$. The sum of (4) then starts with the pair terms, u_{aa} and u_{ab}. These in turn will be functions only of the distances, r_{ij}, in this case. The approximation of including only the pair potentials is then that of setting $u_3 \equiv 0$ for all sets of 3 molecules.

If the summation of eq. (4) is extended to functions of triplets, quadruplets, and higher multiplets, it is not an approximation, but defines the u_n implicitly. The explicit expression is

$$u_n\{\boldsymbol{n}\} = \sum_{\{\nu\}} (-1)^{n-\nu} U_\nu\{\boldsymbol{\nu}\}_n. \qquad (6.1.5)$$

To prove this insert (5) in (4), to obtain

$$\begin{aligned} U_N\{N\} &= \sum_{\{n\}_N} \sum_{\{\nu\}_n} (-)^{N-\nu} U_\nu\{\boldsymbol{\nu}\}_n \\ &= \sum_{\{\nu\}_N} [\sum_{\{m\}_{N-\nu}} (-)^m] U_\nu\{\boldsymbol{\nu}\}_N. \end{aligned} \qquad (6.1.6)$$

For every $\{\boldsymbol{\nu}\}_N \not\equiv \{N\}$ the sum over $\{\boldsymbol{m}\}_{n-\nu}$ includes all subsets, $\{\boldsymbol{m}\}_{N-\nu}$ that are members of $\{N - \boldsymbol{\nu}\}_N$, and since the number of these for each numerical value of $m = m_a, m_b, \ldots$ is given by (3) we have

$$\begin{aligned} \sum_m \frac{(N - \boldsymbol{\nu})!}{(N - \boldsymbol{\nu} - \boldsymbol{m})!\,\boldsymbol{m}!} (-)^m &= \prod_a \frac{(N_a - \nu_a)!\,(-)^{m_a}}{(N_a - \nu_a - m_a)!\,m_a!} \\ = \prod_a (1 - 1)^{N_a - \nu_a} &= \prod_a \delta(N_a - \nu_a), \end{aligned}$$

namely zero unless $\nu_a = N_a$ for all species. The one remaining term is the identity.

If the forces between molecules were independent of the internal coordinates, the functions $u_2, u_3, \ldots$ would be functions of $\mathbf{r}_{ia}, \mathbf{r}_{jb}, \ldots$ alone. The Hamiltonian would be separable in the internal coordinates. Use the complete expression (1.4.2) for the grand canonical partition function, sum over the internal quantum states of each of the N_a molecules of species a for each a, and, with the classical limit for the energy as a function of the centers of mass, $\mathbf{r}_{ia}$, and their conjugate momenta, $\mathbf{p}_{ia}$, integrate over the momenta. One obtains the factor $\lambda_a^{-3}Q_{ia}$ for each molecule of species a. Define an activity

$$z_a = (\lambda_a^{-3}Q_{ia}) \exp(\mu_a/kT), \tag{6.1.7}$$

and one has, from (1.4.3),

$$e^{PV/kT} = \sum_N \frac{\mathbf{z}^N}{\mathbf{N}!} \iint_V \cdots \int [\exp - (U_N\{N\}/kT)] d\{N\}, \tag{6.1.8}$$

where

$$\mathbf{z}^N = \mathbf{z}_a^{N_a} z_b^{N_b} \ldots, \tag{6.1.9}$$

and the potential energy $U_N\{N\}$, depends only on the coordinates of the centers of mass, $\{N\} = \mathbf{r}_{1a} \ldots \mathbf{r}_{Na}, \mathbf{r}_{1b}, \ldots$. As before we see that in the perfect gas limit for which $U_N \equiv 0$ the integrals are V^N and one has $PV = kT\ V \sum_a z_a$. The activities, z_a, of (7) are normalized so that $\rho_a = \partial(P/kT)/\partial \ln z_a = z_a$ in the perfect gas limit. Instead of (7) we could use

$$z_a = \lim_{\rho_a \to 0} \{\rho_{a0} \exp[\mu_a - \mu_a(\rho_{a0})]/kT\}, \tag{6.1.10}$$

as the definition of the activity, z_a.

Now the approximation that the pair potentials between different molecules are independent of their internal coordinates is probably fairly good for molecules without dipole moments, and which are not long chain organic molecules. This is pragmatically justified by the fact that in the fluid state, at least, such molecules appear to obey the same law of corresponding states as the truly spherical molecules such as argon (section 6). In any case one can formally avoid this assumption. We may simply sum exp $- E_K/kT$ over all internal quantum states at fixed values of the coordinates of the centers of mass, then integrate over the center of mass momenta, multiply by (λ_a^3/Q_{ia}) for each molecule of species a, and define the resulting function of the coordinates of the centers of mass as exp $-U_N^*\{N\}/kT$. An "average" potential U_N^* is thus defined. This average potential, U_N^*, may, it is true, then be a function of T, which does not alter the manipulations of this chapter. Stated in more erudite quantum mechanical language we define $U^*_N\{N\}$ as $-kT$ times the logarithm of the diagonal portion of the partial trace of the coordinate representation of the matrix $\prod_a (\lambda_a{}^3/Q_{ia})^{N_a}$ exp $-H/kT$, the partial trace being taken over the internal coordinates.

This sort of formalism permits us to regard the treatment in terms of the center of mass coordinates alone as completely general, but it may not always be the method of choice. The reason is as follows.

By defining the component potentials, u_n with equation (6.1.4) or (6.1.5), we obviate the necessity of assuming at the outset that the total potential energy is a sum of pair terms, and will be able to carry through a formal development independent of such an assumption. At the end, however, in order to use any equations numerically we must return to at least the assumption that the component potentials, u_n, become numerically negligible as n increases much beyond two, and usually actually assume $u_n = 0$ for $n > 2$. Although certainly not exact, this assumption probably leads to very good results for molecules that do not react with each other to form true chemical bonds. (i.e. it is completely invalid for H atoms, or for free radicals such as CH_3). However, if the molecules have large dipole moments, so that the pair potentials depend strongly on the angular orientation variables, or if we deal with long chain polymers, then the sum of pair potential assumption is only nearly valid if the significant internal coordinates are retained. Actually no specific change in language is needed to include the case that internal coordinates are present, but in the interest of simplicity we shall carry through most of the equations as though internal coordinates were absent.

6.2. Distribution functions

The number density functions, $\rho_n\{\mathbf{n}\}$, are defined by the statement,

> $\rho_n\{\mathbf{n}\}\, d\{\mathbf{n}\} \equiv$ the probability that at the coordinate position $\{\mathbf{n}\}$ within the volume element $d\{\mathbf{n}\}$, there are simultaneously at all positions, molecules of the designated species. (6.2.1)

Thus if $n = 1$, $\mathbf{n} =$ one molecule of species a, $\rho_a(\mathbf{q}_a)$ is the probability density of finding a molecule of species a at $\mathbf{q}_a$. If $\mathbf{q}_a = \mathbf{r}_a$ only, then in a fluid $\rho_a(\mathbf{r}_a)$ will be independent of the coordinate and equal to

ρ_a, the average number density of species a. In general

$$\lim_{V\to\infty} [V^{-1}\int\rho_a(\boldsymbol{q}_a)J(q_a)d\boldsymbol{q}_a] = \rho_a, \tag{6.2.2}$$

although for single crystals $\rho_a(\boldsymbol{r}_a)$ will be peaked at the lattice positions. For a fluid the integral over the internal part of $\boldsymbol{q}_a$ of $J(\boldsymbol{q}_a)\,d\boldsymbol{q}_a$ (internal) always gives ρ_a.

For a pair, $\rho_{ab}(\boldsymbol{q}_a,\boldsymbol{q}_b)$ gives the probability density of observing simultaneously a molecule of species a at $\boldsymbol{q}_a$, and one of b at $\boldsymbol{q}_b$. For a single thermodynamic phase (and we shall always tacitly assume this restriction) we have that as $\boldsymbol{q}_a$ and $\boldsymbol{q}_b$ are widely separated,

$$\lim_{r_{ab}\to\infty} [\rho_{ab}(\boldsymbol{q}_a,\boldsymbol{q}_b)] = \rho_a(\boldsymbol{q}_a)\rho_b(\boldsymbol{q}_b), \tag{6.2.3}$$

since the presence or absence of the molecule of species a at $\boldsymbol{q}_a$ does not affect the probability that b be at $\boldsymbol{q}_b$ when the distance is very large. In a fluid, if $\boldsymbol{q}_a = \boldsymbol{r}_a$, $\boldsymbol{q}_b = \boldsymbol{r}_b$, that is if we do not use the internal coordinates, ρ_{ab} will depend on only the distance, r_{ab} between then. In a one-component system, $\rho_2(r) = \rho_{aa}(r_{aa})$ is experimentally measurable as the three-dimensional Fourier transform of the X-ray or neutron scattering pattern. Its general form is given in Fig. 6.1.

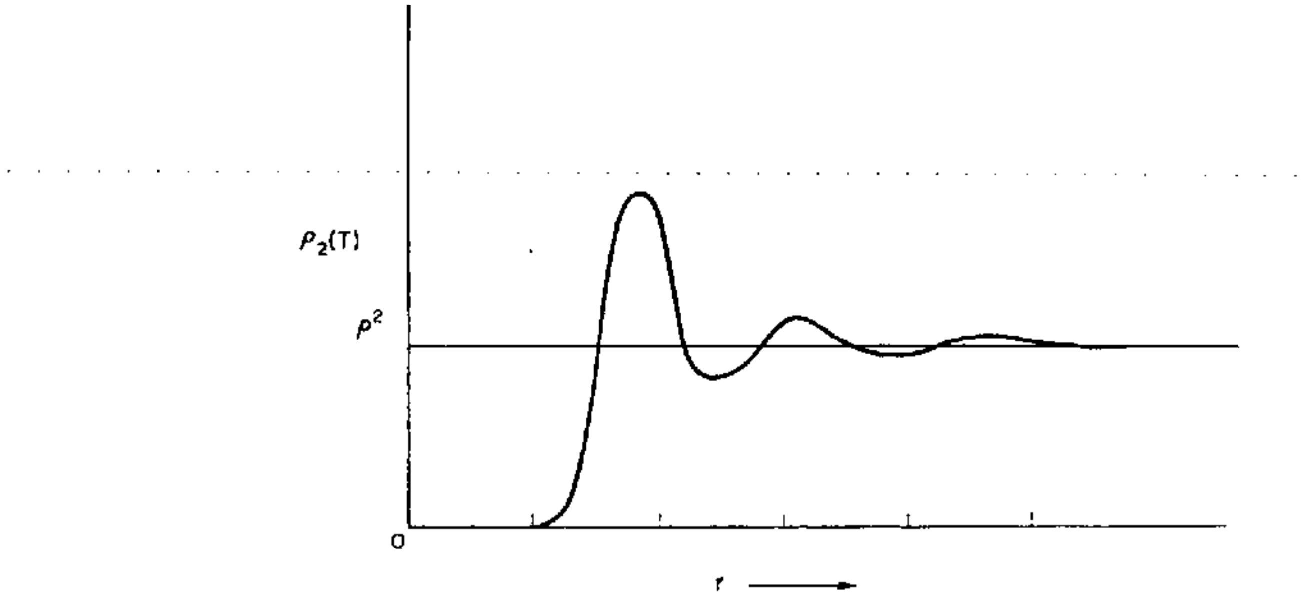

FIG. 6.1. Schematic representation of $\rho_2(r)$ for a liquid.

Again there is, as there so often is, a source of confusion due to numbering. In a crystal, the probability density $\rho_{ab}(\boldsymbol{r}_1,\boldsymbol{r}_2)$ of finding a molecule of species a at $\boldsymbol{r}_1$, and one of species b at $\boldsymbol{r}_2$ is *not* generally equal to the probability density $\rho_{ba}(\boldsymbol{r}_1,\boldsymbol{r}_2)$ of finding a molecule of species b at $\boldsymbol{r}_1$ and of species a at $\boldsymbol{r}_2$. In a fluid, since $|\boldsymbol{r}_1 - \boldsymbol{r}_2| = |\boldsymbol{r}_2 - \boldsymbol{r}_1|$ and

ρ_{ab} depends only on $r_{12} = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$ we have

$$\rho_{ab}(r_{12}) = \rho_{ba}(r_{12}).$$

We shall retain the normalization appropriate to the crystalline case that $\rho_{ab}(r_{12})$ is the probability density of finding a molecule of species a at some fixed position, $\boldsymbol{r}_1$, and simultaneously of finding one of type b at another position $\boldsymbol{r}_2$ a distance r_{12} away from the position $\boldsymbol{r}_1$. This is half the probability of finding a pair of molecules, one of type a, the other of type b. We then have in a system of σ components that

$$\rho(r) = \sum_{a \geqslant 1}^{a=\sigma} \sum_{b \geqslant 1}^{b=\sigma} \rho_{ab}(r) \tag{6.2.4}$$

gives the total number density of finding a molecule, of whatever species, at two positions a distance r apart. Were the other normalization used, $\rho_{ab}{}^* = 2\rho_{ab}$, we would have to write

$$\rho = \sum_{\sigma \geqslant a \geqslant b \geqslant 1} \sum \rho_{ab}{}^* \tag{6.2.4$'$}$$

so that the unlike terms, $ab = ba$ occurred only once in the summation.

It is somewhat more convenient to use the distribution functions, $F_{\boldsymbol{n}}\{\boldsymbol{n}\}$, instead of the number density functions,

$$\rho_n\{\boldsymbol{n}\} = \rho_n F_n\{\boldsymbol{n}\}, \tag{6.2.5}$$

$$\boldsymbol{\rho}^{\boldsymbol{n}} = \rho_a{}^{n_a}\rho_b{}^{n_b} \cdot \cdot \cdot,$$

since then the normalization is always the same, namely,

$$\lim_{V \to \infty} \left[V^{-n} \int\int_V \cdot \cdot \cdot \int F_n\{\boldsymbol{n}\}\, d\boldsymbol{n} \right] = 1. \tag{6.2.6}$$

In the fluid, then, $F_a \equiv 1$ if the coordinate is only $\boldsymbol{r}_a$, and its integral over the internal coordinate space is always unity.

We can generalize eq. (3) for any number set $\boldsymbol{n}$. If the number set is written as a sum, $n + m$, and the distance $r_{n,m}$ between every member of the set n and every member of the set $\boldsymbol{m}$ becomes very large, then

$$\lim_{r_{nm} \to \infty} [F_{n+m}(\{\boldsymbol{n} + \boldsymbol{m}\})] = F_n\{\boldsymbol{n}\}F_m\{\boldsymbol{m}\}. \tag{6.2.7}$$

This property of the functions is one that we shall use repeatedly; it is, however, subject to a certain limitation that is discussed in section 6.11.

6.3. Equations for the distribution functions

The reduced probability density, $\rho^n F_m\{m\}$, can obviously be obtained by summing the probabilities, $[\exp(-PV/kT)](z^N/N!)\exp(-U/kT)$, of finding the set N of molecules in V, over all $N > n$, and integrating over all positions of the molecules other than those at $\{n\}$. A little care is needed, however, to decide the weighting to give each term of N. Write $N = n + M$. We must then sum over M from zero to infinity, and we integrate over $d\{M\}$, keeping the coordinates of the numbered set n fixed at $\{n\}$. However, we could have chosen the numbered set which remains at $\{n\}$ in any of $(n + M)!/n!M!$ ways. But we defined $\rho_n\{n\}$ as being the probability density that *any* molecule of the correct species be at $\{n\}$, not a particular numbered molecule $1a$ at position q_{1a}, $2b$ at $q_{2b}, \ldots$, etc. We must therefore multiply by $n!$. This means we multiply the term M by $N!/M!$ We then multiply both sides by $z^{-n} \exp PV/kT$ to obtain

$$e^{PV/kT}\left(\frac{\rho}{z}\right)^n F_n\{n\} = \sum_{M\geqslant 0} \frac{z^M}{M!} \int\!\!\int_V \cdots \int \{\exp[-U_{n+M}\{n+M\}/kT]\} d\{M\}. \tag{6.3.1}$$

If we arbitrarily define

$$F_0 \equiv 1, \tag{6.3.2}$$

then this equation becomes (6.1.8) for $e^{PV/kT}$ when $n \equiv 0$.

The distribution functions, F_n, depend on the intensive properties, z and T, of the ensemble, $F_n = F_n(z,T\{n\})$. In the limit that all activities, z_a, approach zero, $\rho_a/z_a = 1$, and $P \to 0$, $\exp PV/kT \to 1$. The sum on the right has then only one non-zero term, $M \equiv 0$, so that

$$F_n(z = 0,T,\{n\}) = \exp - (U_n\{n\}/kT, \tag{6.3.3}$$

the distribution functions of zero activity are $\exp - U/kT$.

An inevitable confusion arises from the mathematical necessity of labeling the coordinates of identical molecules with indices to distinguish the coordinates from each other. The probability, W_N, of finding the set N of molecules in V is,

$$W_N = e^{-PV/kT} \frac{z^N}{N!} \int\!\!\int_V \cdots \int e^{-U_N/kT} d\{N\}. \tag{6.3.4}$$

From (6.1.8) one then has, as one must, that

$$\sum_{N\geqslant 0} W_N = 1. \tag{6.3.5}$$

Now one may logically define two differently normalized functions, $W_N\{N\}$, or $W_N{}^*\{N\}$, as probability density functions of having exactly a set, N, of molecules in V and at the coordinate positions $\{N\}$. We prefer to define

$$W_N\{N\} = \mathbf{z}^N e^{-PV/kT} e^{-U_N\{N\}/kT}. \tag{6.3.6}$$

as the probability density that the (necessarily) numbered coordinate positions $\{N\}$ are occupied by molecules of the appropriate species, molecules of type a at the positions $\{N_a\}$, etc. We then have

$$W_N = \iint_V \cdots \int \frac{W_N\{N\}}{N!} d\{N\}, \tag{6.3.7}$$

since in the integration the same physical configuration is covered $N!$ times.

We could alternatively have defined

$$W_N{}^*\{N\} = (\mathbf{z}^N/N!)\, e^{-PV/kT} e^{-U_N\{N\}/kT} \tag{6.3.8}$$

as the probability that molecule $1a$ was at position $1a$, molecule $2a$ at position $2a$, ... We then have

$$W_N = \iint_V \cdots \int W_N{}^*\{N\} d\{N\}, \tag{6.3.9}$$

which appears more convenient than (7), but the probability of observing necessarily indistinguishable molecules at positions $\{N\}$ is

$$W_N\{N\} = N!\, W_N{}^*\{N\}. \tag{6.3.10}$$

Our probability density functions $\rho_n\{n\} = \boldsymbol{\rho}^n F_n\{n\}$ are defined analogously to W_N of (7), namely as the probability densities that molecules of the appropriate types are at $\{n\}$. This seems to be the only normalization that one could use in an open system, since if the system is exposed to an infinite sea of molecules it has no meaning to ask the probability of finding some prenumbered molecule $1a$ at position $1a$, . . ., etc. Even in a closed system having exactly N_a molecules of kind a in V the normalization would be awkward, although logical, and the probability density, $\rho_n\dagger\{n\}$, of finding molecule $1a$ at position $1a$, $1b$ at position $1b$, . . ., etc. would be

$$\rho_n\dagger\{n\} = p_n\{n\}/N!, \tag{6.3.11}$$

so that the values at any position $\{n\}$ would depend critically on the size of the system as well as on the intensive variables z and T.

Our normalization, (6.2.5), of $F_n\{n\}$ is such that if *all* distances r_{ij} become large, and if the coordinates contain no internal coordinates, then F_n approaches unity in value for a fluid. With (4) this means that for sufficiently large V, and finite n†

$$\iint_{\Gamma} \cdots \int \rho_n\{n\}\, d\{n\} \to \prod_a \bar{N}_a{}^{n_a} = \bar{N}^n \prod_a x_a{}^{n_a}. \tag{6.3.12}$$

The integral of ρ_{ab} for $a = b$ is then $\bar{N}_a\bar{N}_b$ in sufficiently large volume V, which is the total number of a–b pairs, and the integral of the two terms, ρ_{ab} plus ρ_{ba}, in the summation of (6.2.4) is twice the number of a–b pairs. The integral of the single term ρ_{aa} in (6.2.4) over V is $\bar{N}_a{}^2$ which is twice the total number of a–a pairs.

If the potential energy is a sum of pair terms only, and we define quantities of the dimensions of energy,

$$\bar{u}_{ba} = \bar{u}_{ab} = \rho \int_0^\infty 4\pi r^2\, u_{ab}(r) F_{ab}(r)\, dr, \tag{6.3.13}$$

so that with $x_a = \rho_a/\rho$,

$$\iint_V u_{ab}(r_{ij})\rho_{ab}(r_{ij}) d\mathbf{r}_i\, d\mathbf{r}_j = \bar{N} x_a x_b \bar{u}_{ab}. \tag{6.3.14}$$

Define the average potential energy $\bar{u}$ per molecule by

$$\bar{u} = \tfrac{1}{2} \sum_{a=1}^{a=\sigma} \sum_{b=1}^{b=\sigma} x_a x_b\, \bar{u}_{ab}, \tag{6.3.15}$$

and the average potential energy of the molecules in V will be

$$\bar{U} = \bar{N}\bar{u}. \tag{6.3.16}$$

We shall now show a simple physical interpretation, in a classical system, of the quantity $-kT \ln F_n$.

6.4. The potentials of average force

An energy function, $\Omega_n(z,T,\{n\})$, can be defined by

$$\Omega_n(z,T,\{n\}) = -kT \ln F_n\,(z,T,\{n\}). \tag{6.4.1}$$

We shall show that, if the system is classical, these functions are potentials of average force.

The meaning of this statement follows. Suppose in a system the set $\mathbf{n}$

---

† In section 6.11 we discuss the case of finite V.

of molecules is at the position $\{\mathbf{n}\}$. Acting on any one of these molecules, say molecule ia, there will be some force along, say, the x-axis, f_{xia}. This force will be composed of additive components, one part due directly to the other molecules of the set $\mathbf{n}$ located at $\{\mathbf{n}\}$, and the other due to all other molecules in the neighborhood of ia, whose average positions will be influenced by the set $\mathbf{n}$ at $\{\mathbf{n}\}$. The force f_{xia} will fluctuate in time, even were the other molecules of $\mathbf{n}$ to remain at $\{\mathbf{n}\}$. Looked at differently, at any time t, the force would be different in different members of the ensemble for which a set $\mathbf{n}$ were at the position $\{\mathbf{n}\}$. There will be however, an ensemble average value $\bar{f}_{zia}$. The assertion is that

$$\bar{f}_{xia}(\mathbf{z},T,\{\mathbf{n}\}) = -\partial\Omega_{\mathbf{n}}(\mathbf{z},T,\{\mathbf{n}\})\partial x_{ia}. \tag{6.4.2}$$

The proof from (6.3.1) is extremely easy. Differentiate (6.3.1) with respect to x_{ia}, on both sides, remembering that this coordinate is one in $\{\mathbf{n}\}$, and not one over which the integration of the left extends. The only part of the left hand side of (6.3.1) that depends on x_{ia} is F_n, and the only part on the right is $U_{\mathbf{n}+\mathbf{M}}$. The differentiation can be made under the integral. Using (1) and (2) and multiplying both sides by $[\exp - PV/kT](\mathbf{z}/\boldsymbol{\rho})^n\, kTF_n{}^{-1}$ one finds, with (2) for $\bar{f}_{xia}$,

$$\bar{f}_{xia} = \sum \iint \cdots \int -\frac{\partial U_{\mathbf{n}+\mathbf{M}}}{\partial x_{ia}}\, W_{\mathbf{n},m}(\{n\}\{M\})\, d\{M\}, \tag{6.4.3}$$

where

$$W_{\mathbf{n},\mathbf{M}} = e^{-PV/kT}\boldsymbol{\rho}^{-\mathbf{n}}\mathbf{z}^{\mathbf{n}+\mathbf{M}}(\mathbf{M}!)^{-1}\, F_{\mathbf{n}}{}^{-1} \exp - U_{\mathbf{n}+\mathbf{M}}/kT. \tag{6.4.4}$$

This is exactly the normalized probability density that if there is a set $\mathbf{n}$ of molecules at $\{\mathbf{n}\}$, there are $\mathbf{M}$ other molecules in V, at positions $\{\mathbf{M}\}$. The normalization is just such that the sum over $\mathbf{M}$ of the integral over $d\mathbf{M}$ of $W_{\mathbf{n},\mathbf{M}}$ is unity. The expression (3) is then exactly the average force along x_{ia} when the set $\mathbf{n}$ is at $\{\mathbf{n}\}$.

6.5. The Kirkwood superposition assumption

Just as we wrote, in equation (6.1.4) the total potential energy function $U_{\mathbf{N}}\{\mathbf{N}\}$ as a sum of component potentials, $u_{\mathbf{n}}\{\mathbf{n}\}_{\mathbf{N}}$, we may always write,

$$\Omega_{\mathbf{n}}\{\mathbf{n}\} = \sum_{\{\boldsymbol{\nu}\}_{\mathbf{n}}} \omega_{\boldsymbol{\nu}}\{\boldsymbol{\nu}\}_{\mathbf{n}}, \tag{6.5.1}$$

$$\omega_{\boldsymbol{\nu}}\{\boldsymbol{\nu}\} = \sum_{\{\boldsymbol{\mu}\}_{\boldsymbol{\nu}}} (-)^{\boldsymbol{\nu}-\boldsymbol{\mu}}\Omega_{\boldsymbol{\mu}}\{\boldsymbol{\mu}\}_{\boldsymbol{\nu}}. \tag{6.5.2}$$

An assumption, introduced by Kirkwood and called by him the assumption of superposition, is that $\omega_\nu = 0$ for $\nu > 2$. Since for fluids, and if no internal coordinates appear, $\Omega_1(r) = \omega_1(r) = 0$, there remains only ω_2, i.e. ω_{aa} or ω_{ab} if different species a, b, are present, and this function will depend only on the distance r_{ij} of the two molecules. In this approximation the higher order distribution functions, F_n, are the products of the pair distribution functions.

Unfortunately there is little evidence, at present, justifying this assumption. Probably the best evidence is very pragmatic, namely that with its use it is possible to solve for F_2 as the solution of an integral equation (section 6.12). The results obtained by Kirkwood and co-workers appear to agree reasonably well with experimental values obtained from X-ray scattering, and also with values obtained for hard spheres by numerical electronic computers.

It is clear that the assumption $\omega_3 \equiv 0$ is incorrect, even if u_3 of (6.1.4) were strictly zero. In one other sense, the assumption $\omega_\nu = 0$, $\nu > 2$ is probably much poorer, than the equivalent assumption for u_ν.

From the second virial coefficient for argon and its temperature dependence one may evaluate the pair potential function $u_2(r)$. This may then be used, with the assumption of pair additivity, to compute the properties of crystalline argon, where the u_n's for $n > 2$ should be important if they differ very significantly from zero. The results agree reasonably with experimental values, but appear to differ outside of experimental error.

Now on the other hand the potential of average force, $\Omega_N\{N\}$, for $N = \bar{N}$ the average number of molecules in V, for values of the coordinates $\{N\}$ all within V, must go over asymptotically to the potential energy function, $U_N\{N\}$. This must be strictly true for a closed system containing exactly the set N. The pair potentials of average force, $\Omega_2(r)$, look very different from the isolated pair potentials, $u_2(r)$, and it is unlikely that the sum of the $\omega_2(r_{ij})$ adequately reproduces U_N. This means that in the sum up to very large values of N, we can hardly expect the sum of ω_ν for $\nu > 2$ to be negligible. This does not negate the possibility that Ω_3 is well given by the sum of the three pair terms ω_2.

Although our formal manipulations will not involve the Kirkwood assumption, all equations from which F_n and hence Ω_n can be evaluated involve F_{n+1} and hence Ω_{n+1}. It is only by use of the Kirkwood assumption, or some similar "closure" that expresses Ω_{n+1} in terms of Ω_n that an evaluation can be made numerically.

6.6. The virial of Clausius

An extremely valuable expression for the pressure in any system gives PV as the perfect gas kinetic energy term $\bar{N}kT$ plus the virial of Clausius which contains the contribution of the mutual forces between molecules.

Use (6.1.8) for $e^{PV/kT}$ as a sum over sets N of integrals over the coordinate space. In this we use dimensionless coordinates, $\boldsymbol{R}_{ia}$, for molecule ia of type a, defined in terms of the usual metric vector $\boldsymbol{r}_{ia}$ by

$$\begin{aligned} \boldsymbol{r}_{ia} &= V^{1/3}\,\boldsymbol{R}_{ia}, \\ d\boldsymbol{r}_{ia} &= V\,d\boldsymbol{R}_{ia}, \end{aligned} \tag{6.6.1}$$

and $\{N'\}$ for the set of these dimensionless coordinates for a set N of molecules,

$$\begin{aligned} \{N'\} &= \boldsymbol{R}_{1a}, \ldots, \boldsymbol{R}_{Na}, \boldsymbol{R}_{1b}, \ldots, \\ V^N d\{N'\} &= V^N d\boldsymbol{R}_{1a} \ldots d\boldsymbol{R}_{Na} d\boldsymbol{R}_{1b} \ldots = d\{N\}. \end{aligned} \tag{6.6.2}$$

The integration range of $d\{N'\}$ is now over unit volume. Equation (6.1.8) is now

$$e^{PV/kT} = \sum_{N\geqslant 0} \frac{(V\boldsymbol{z})^N}{N!} \iint_1 \ldots \int [\exp - U_N(V^{1/3}\{N'\})/kT] d\{N'\}. \tag{6.6.3}$$

Carry out the operation

$$\boldsymbol{O} = VkT\,e^{-PV/kT}(\partial/\partial V)_{\boldsymbol{z},T} \tag{6.6.4}$$

on both sides of this equation.

Since the pressure P depends only on $\boldsymbol{z}$ and T, and not on the total volume, the operator $\boldsymbol{O}$ of (4) operating on the left hand side of (3) gives,

$$\boldsymbol{O}e^{PV/kT} = PV. \tag{6.6.5}$$

On the right hand side $\boldsymbol{O}$ acts on $(V\boldsymbol{z})^N$ and on the potential energy function. We have

$$\boldsymbol{O}(V\boldsymbol{z})^N = kTN(V\boldsymbol{z})^N e^{-PV/kT}. \tag{6.6.6}$$

The term

$$\begin{aligned} &\frac{(V\boldsymbol{z})^N}{N!} e^{-PV/kT} \iint_1 \ldots \int [\exp - U_N(V^{1/3}\{N'\})/kT] d\{N'\} \\ &= \frac{\boldsymbol{z}^N}{N!} e^{-PV/kT} \iint_V \ldots \int [\exp - U_N\{N\}/kT] d\{N\} \\ &= W_N \end{aligned} \tag{6.6.7}$$

is the probability of finding the set N of molecules in the open system, and the sum of (6.6.6) over N yields

$$kT \sum_{N \geqslant 0} N W_N = \bar{N}kT. \tag{6.6.8}$$

If the potential energy were identically zero this would be the only term on the right and we would have the perfect gas expression, $PV = \bar{N}kT$. The additive term due to the operator $\boldsymbol{O}$ acting on $\exp - U_N/kT$ under the sum and integral is the Clausius virial $- \frac{1}{3} < U' >$, and

$$PV - \bar{N}kT = \tfrac{1}{3} < U' >. \tag{6.6.9}$$

To evaluate $<U'>$ we note that the potential energy function, U_N, of (3) is a function of the $3N$ metric vector components, $x_{i\alpha\alpha} = V^{1/3}X_{i\alpha\alpha}$ of $\alpha = 1, 2, 3$ and $1a < ia < N_a$, for all types, a, of molecules. The logarithmic derivative, $V(\partial/\partial V)$, operating on U_N gives

$$V(\partial/\partial V)U_N = \tfrac{1}{3} \sum_{i\alpha\alpha=1}^{3N} \left(x_{i\alpha\alpha} \frac{\partial U_N}{\partial x_{i\alpha\alpha}} \right) = \tfrac{1}{3} U'_N\{N\}, \tag{6.6.10}$$

and

$$\boldsymbol{O}e^{-U_N/kT} = - \tfrac{1}{3} e^{-PV/kT} e^{-U_N/kT} U'_N. \tag{6.6.11}$$

The virial $- \frac{1}{3} <U'>$ is then

$$\begin{aligned} -\tfrac{1}{3} <U'> &= -\tfrac{1}{3} \sum_{N \geqslant 0} \frac{z^N}{N!} \iint_V \cdots \int e^{-(PV+U_N)/kT} U'_N \, d\{N\} \\ &= -\tfrac{1}{3} \sum_{N \geqslant 0} \frac{1}{N!} \iint_V \cdots \int W_N\{N\} U'_N\{N\} \, d\{N\}, \end{aligned} \tag{6.6.12}$$

which is, as our notation has anticipated, minus one third of the average value of the function U'_N.

The expression for U'_N can be put in a simple form if we assume that the potential energy function is a sum of pair terms only. The pair term, $u_{ij}(r_{ij})$, is a function of the distance, r_{ij} between the two molecules i and j. Since

$$x_{i\alpha} \partial r_{ij} / \partial x_{i\alpha} = x_{i\alpha}(x_{i\alpha} - x_{j\alpha})/r_{ij},$$

$$x_{j\alpha} \partial r_{ij} / \partial x_{j\alpha} = - x_{j\alpha}(x_{i\alpha} - x_{j\alpha})/r_{ij},$$

$$\sum_{\alpha=1}^{\alpha=3} [x_{i\alpha}(\partial/\partial x_{i\alpha}) + x_{j\alpha}(\partial/\partial x_{j\alpha})] u_{ij}(r_{ij})$$

$$= \sum_{\alpha=1}^{\alpha=3} \frac{(x_{i\alpha} - x_{j\alpha})^2}{r_{ij}} \frac{du_{ij}(r_{ij})}{dr_{ij}} = r_{ij} \frac{du_{ij}(r_{ij})}{dr_{ij}}. \tag{6.6.13}$$

We have, then,

$$U'_N = \sum_{N \geqslant i > j \geqslant 1} \sum r_{ij} du_{ij}(r_{ij})/dr_{ij}. \tag{6.6.14}$$

We may use the pair probability densities, $\rho_{ab}(r) = \rho_a \rho_b F_{ab}(r)$ to evaluate the averages. Analogously to 6.3.13 we define

$$\bar{u}'_{ab} = \rho \int_0^\infty 4\pi r^2 \left[r \frac{du_{ab}(r)}{dr} \right] F_{ab}(r) dr, \tag{6.6.15}$$

and

$$\bar{u}' = \tfrac{1}{2} \sum_{a=1}^{a=\sigma} \sum_{b=1}^{b=\sigma} x_a x_b \bar{u}'_{ab}. \tag{6.6.16}$$

We then have, from (9), since

$$<U'> = \bar{N}\bar{u}', \tag{6.6.17}$$

that

$$P = \rho[kT - \tfrac{1}{3}\bar{u}']. \tag{6.6.18}$$

6.7. Calculation of $F_n(\mathbf{z},T,\{\mathbf{n}\})$ from $F_n(\mathbf{y},T,\{\mathbf{n}\})$

In this section we shall derive equations by which the distribution functions, $F_n(\mathbf{z},T,\{\mathbf{n}\})$ are given in terms of the functions for any other activity set $\mathbf{y} = y_a, y_b \ldots$. In section 6.3, eq. (6.3.3) we saw that $[\exp - U_N\{N\}/kT]$ could be written as the limit, when all $z_a \to 0$, of F_N. Equation (6.3.1) thus gives $F_n(\mathbf{z})$ in terms of $F_N(\mathbf{z} = 0)$. We shall show how $F_n(\mathbf{z})$ can be obtained from $F_N(\mathbf{y})$ for any $\mathbf{y}$.

For the sake of brevity we define

$$G_n(\mathbf{z},T,\{\mathbf{n}\}) = \frac{\rho(\mathbf{z},T)^n}{\mathbf{z}} \exp \frac{VP(\mathbf{z},T)}{kT} F_n(\mathbf{z},T,\{\mathbf{n}\}). \tag{6.7.1}$$

Since at $\mathbf{z} \equiv 0$ we have $P = 0$, and $\rho_a/z_a = 1$ for all a, one has that

$$G_n(\mathbf{z} \equiv 0,T,\{\mathbf{n}\}) = F_n(\mathbf{z} \equiv 0,T,\{\mathbf{n}\}) = \exp - U_n\{\mathbf{n}\}/kT, \tag{6.7.2}$$

and (6.3.1) becomes

$$G_n(\mathbf{z},t,\{\mathbf{n}\}) = \sum_{\mathbf{M} \geqslant 0} \frac{\mathbf{z}^{\mathbf{M}}}{\mathbf{M}!} \int\int_V \cdots \int G_{n+M}(\mathbf{z} \equiv 0,T,\{\mathbf{n}+\mathbf{M}\}) d\{\mathbf{M}\}. \tag{6.7.3}$$

Now (3) represents an infinite set of equations for all values $0 \leqslant \mathbf{n} \leqslant \infty$, determining G_n for $\mathbf{z}$ in terms of the G_N for $\mathbf{z} \equiv 0$.

Actually, for all real systems, U_N becomes very large and positive if any pair distance, r_{ij}, approaches zero. We would not change the answers if we assumed $U \to \infty$ if any pair distance $r_{ij} < r_0$ where r_0 is suitably chosen. In this case all functions G_N would be identically zero for all values of the coordinates if N were larger than some N_0, $N_0 = V/(\pi r_0^3/6)$. We would then have a finite number of (non-zero) functions $G_N(z)$ and an equal finite number, $G_N(z \equiv 0)$, and just this number of equations. As a mathematical exercise we might choose to regard the functions $G_N(z)$ as known, those for $G_N(z \equiv 0)$ as unknown, but determined by the set of integral equations (3), in terms of $G_N(z)$. We would hardly expect to be able to solve this set explicitly. It so happens that we can actually write the solution by inspection, and readily prove its correctness.

The solution is,

$$G_n(z \equiv 0,T,\{n\}) = \sum_{N\geqslant 0} \frac{(-z)^N}{N!} \iint_V \cdots \int G_{n+N}(z,T,\{n + N\}\, d\{N\}. \tag{6.7.4}$$

To prove (4) correct insert it in (3). We have

$$G_n(z,T,\{n\}) = \sum_{N\geqslant 0} \sum_{M\geqslant 0} \frac{z^M(-z)^N}{M!N!} \iint_V \cdots \int G_{n+M+N} \times (z,t,\{m + M + N\}dMdN. \tag{6.7.5}$$

In this write $N + M = K$ and $M = K - N$. One has $\int dMdN = \int dK$. The coefficient of the term G_{n+K} is

$$\sum_{N=0}^{N=K} \frac{z^{K-N}(-z)^N}{(K-N)!N!} = \prod_a \sum_{N_a=0}^{N_a=K_a} \frac{z_a^{K_a-N_a}(-z_a)^{N_a}}{(K_a - N_a)!N_a!}$$

$$= \prod_a \frac{(z_a - z_a)^{K_a}}{K_a!} = \prod_a \delta(K_a), \tag{6.7.6}$$

namely zero unless all K_a are zero, and if all $K_a = 0$ the coefficient is unity. An identity in (5) is obtained. Thus (4) is proven correct.

Now we can use (4) to obtain an equation for $G_n(z)$ in terms of the functions $G_n(y)$ for any other activity set y. Namely insert (4) with y replacing z into eq. (3). We obtain an equation similar to (5) with

$G_{n+M+N}(y)$ under the integral and $(-y)^N$ in front. Going through the steps of (6) we find that the coefficient of G_{n+K} is,

$$\prod_a \frac{(z_a - y_a)^{K_a}}{K_a!} = \frac{(z-y)^K}{K!} \tag{6.7.7}$$

so that, replacing K by N, we have

$$G_n(z,T,\{n\}) = \sum_{N\geqslant 0} \frac{(z-y)^N}{N!} \int\int_V \cdots \int G_{n+N}(y,T,\{n+N\})d\{N\}. \tag{6.7.8}$$

In the limit $y \equiv 0$ this equation is the same as eq. (3). If $z \equiv 0$ it is the same as (4) with y replacing z.

Finally, going back to (1) we write (8) in terms of the functions Fn as

$$\begin{aligned}&[\exp P(z,T)V/kT][\rho(z,T)/z]^n F_n(z,T,\{n\})\\ &\quad = \left[\exp \frac{P(y,T)V}{kT}\right]\left[\frac{\rho(y,T)}{y}\right]^n \sum_{N\geqslant 0} \frac{1}{N!}\left[(z-y)\frac{\rho(y,T)}{y}\right]^N\\ &\qquad \int\int_V \cdots \int F_{n+N}(y,T\{n+N\})d\{N\}.\end{aligned} \tag{6.7.9}$$

In the particular case $n \equiv 0$ this is, with $F_0 \equiv 1$,

$$\begin{aligned}\exp\frac{[P(z,T) - P(y,T)^V}{kT} &= \sum_{N\geqslant 0} \frac{1}{N!}\left[(z-y)\frac{\rho(y,T)}{y}\right]^N\\ &\qquad \int\int_V \cdots \int F_N(y,T\{N\})d\{N\}.\end{aligned} \tag{6.7.10}$$

6.8. The osmotic pressure

A particular case of (6.7.10) is one which leads to an expression for the osmotic pressure.

Let one component, say component $a = \alpha$, of the molecular constituents be designated as the solvent, the other components a, b, etc. as solutes. Choose the set y as

$$y = y_\alpha,\ y_a = 0,\ y_b = 0, \ldots \tag{6.8.1}$$

and the set z as

$$z = (z_\alpha = y_\alpha),\ z_a,\ z_b, \ldots \tag{6.8.2}$$

The activity set y is now that which describes the pure solvent, or, more accurately in our language, the solution of a, b, . . . in α in the limit of zero activities or zero concentration of the solute species. The activity set $\mathbf{z}$ is that of the solution with non-zero activities of the solute species, but with the same fixed activity, y_α of the solvent. The difference of pressures,

$$P(\mathbf{z},T) - P(\mathbf{y},T) = P_{\text{osm}}(z_a, z_b, \ . \ . \ . \ T), \tag{6.8.3}$$

is now the osmotic pressure of the solution, the difference in pressure of solution minus that of solvent at the same activity, y_α, of the solvent.

The factor $(\mathbf{z} - \mathbf{y})^{\mathbf{N}}$ under the sum of (6.7.10) contains one factor, $(y_\alpha - y_\alpha)^{N_\alpha}$, which is identically zero unless $N_\alpha = 0$. The sum over $\mathbf{N}$, then, contains only non-zero terms if $N_\alpha = 0$, and can be regarded as a sum over the set,

$$\mathbf{N} = N_a, N_b, \ . \ . \ ., \tag{6.8.4}$$

omitting N_α.

Now examine the quantity

$$\zeta_a = z_a \left\{ \lim_{y_a, y_b \to 0} \left[\frac{\rho_a(y_\alpha, y_a, y_b, \ . \ . \ ., T)}{y_a} \right] \right\}. \tag{6.8.5}$$

The quantity ζ_a is an activity, since it is proportional to z_a, and hence to $e^{\mu_a/kT}$. The proportionality is such that

$$\lim_{\zeta_a, y_b \to 0} \left[\frac{\zeta_a}{\rho_a(y_\alpha, z_a, y_b, \ . \ . \ ., T)} \right] = \lim_{\substack{z_a \to 0, \\ y_a, y_b \to 0}} \left[\frac{z_a \rho_a(y_\alpha, y_a, y_b, \ . \ . \ ., T)}{y_a \rho_a(y_\alpha, z_a, y_b, \ . \ . \ ., T)} \right] \tag{6.8.6}$$

$$= 1.$$

We therefore find that ζ_a is the activity of species a in the solution, normalized so that it becomes equal to the number density, ρ_a of species a in the solution of solvent α at activity y_α. Equation (6.7.10) is then

$$e^{P_{\text{osm}} V/kT} = \sum_{\mathbf{N}} \frac{\boldsymbol{\zeta}^{\mathbf{N}}}{\mathbf{N}!} \int\int_V \ . \ . \ . \int e^{-\Omega_{\mathbf{N}}(\{\mathbf{N}\})/kT} d\{\mathbf{N}\}, \tag{6.8.7}$$

in which the sum runs over only the solute molecule set $\mathbf{N} = N_a, N_b$, and in which the properties of the solvent molecules nowhere explicitly appear!

The properties of the solvent, α, is of course implicitly hidden in the potential of average force function, $\Omega^{\mathbf{N}}$. This is the potential of average force between a set $\mathbf{N}$ of solute molecules in the limiting activity y_α and $y_a \to 0$, $y_b \to 0$ . . .. If the forces between solute molecule pairs a–a,

a–b, b–b, etc., are measured in an infinitely dilute solution of the solvent, which means in the limit that only the pair considered is present in the solvent, and the potentials $\omega_{aa}(r_{aa})$, $\omega_{ab}(r_{ab})$, etc., constructed, then these pair potentials are those that appear in Ω_N. Insofar as the Kirkwood superposition is valid, these alone determine Ω_N, and the computation of the osmotic pressure can be made by exactly the same method as was used in the treatment of the imperfect gas.

In any case the potentials Ω_N that enter (7) are those between the set N of solute species in the pure solvent, and in the absence of any other solute molecules. For sufficiently small values of activity only terms of small values of N are important, and Ω_N is zero over the greater configuration space where all pair distances are large. The integrals are then V^N and one has the perfect solution equation,

$$P_0/kT = \sum_a \zeta_a, \ (\lim \boldsymbol{\zeta} \to 0). \tag{6.8.8}$$

Differentiation yields

$$\rho_a = [\partial(P_0/kT)/\partial \ln \zeta_a]_{T, \zeta_b}, \ldots = \zeta_a \quad (\lim \boldsymbol{\zeta} \to 0) \tag{6.8.9}$$

which agrees with the limit of (6.8.6).

6.9. The cluster development for osmotic pressure

The cluster development of chapter 4 was made with the specific limitation to a pair potential sum for the total potential energy, and for a one-component system. These limitations are unnecessary. Our general notation for the coordinates of sets of molecules permits a fairly easy and concise development in the general case.

We use the notation $\{k\{\boldsymbol{n}_i\}_N\}_u$ to indicate a partition of the set $\boldsymbol{N}$ of numbered molecules into k unconnected subsets $\boldsymbol{n}_i$, $1 \leqslant i \leqslant k$, with coordinates $\{\boldsymbol{n}_i\}_N$, such that

$$\sum_{i=1}^{i=k} \{\boldsymbol{n}_i\}_N = \{\boldsymbol{N}\}, \tag{6.9.1}$$

so that every numbered molecule occurs in one and only one subset $\{\boldsymbol{n}_i\}_N$. Use

$$(\Sigma\{k\{\boldsymbol{n}_i\}_N\}_u) = \text{summation over all partitions,} \tag{6.9.2}$$

including the partition with $k = 1$, $\{\boldsymbol{n}_1\}_N = \{\boldsymbol{N}\}$, to the partition $k = N$ where every subset, $\{\boldsymbol{n}_i\}_N$ contains a single molecule only.

The cluster functions $g_n\{n_i\}_N$ are then defined implicitly by the equation,

$$F_N = (\Sigma\{k\{n_i\}N\}_u) \prod_{i=1}^{i=k} g_{n_i}\{n_i\}_N. \tag{6.9.3}$$

The inverse equation giving g_n explicitly is

$$g_n = (\Sigma\{k\{\nu_i\}_n\}_u)(-)^{k-1}(k-1)! \prod_{i=1}^{i=k} F_{\nu_i}\{\nu_i\}_n. \tag{6.9.4}$$

Define cluster integrals b_n by

$$b_n(V) = \frac{1}{Vn!} \int\int_V \cdots \int g_n\{n\}\, d\{n\}, \tag{6.9.5}$$

$$b_1 = 1 \tag{6.9.6}$$

as in (4.2.6). There are

$$N!/\prod_n (n!)^{m_n} m_n! \text{ ways} \tag{6.9.7}$$

of forming m_n clusters of subsets n out of the set N of molecules, where

$$\Sigma n m_n = N. \tag{6.9.8}$$

The sum of integrals of (6.8.7) is then

$$\sum_N \frac{\zeta^N}{N!} \int\int_V \cdots \int F_N\{N\}\, d\{N\} = \prod_n \sum_{m_n \geqslant 0} \frac{(Vb_n\zeta^n)^{m_n}}{m_n!}$$
$$= \exp V \sum_n b_n \zeta^n, \tag{6.9.9}$$

and

$$P_{\text{osm}}/kT = \sum_n b_n \zeta^n. \tag{6.9.10}$$

Since, in general,

$$[\partial(P_{\text{osm}}/kT)\partial \ln \zeta_a]_{T,y_\alpha,\zeta_b,\ldots} = \rho_a \tag{6.9.11}$$

we have

$$\rho_a = \sum n_a b_n \zeta^n \tag{6.9.12}$$

for the concentration, ρ_a, in molecules per unit volume, of solute species a. One must remember that the b_n's are to be computed from potentials of average force between infinitely dilute solute molecules *at the same activity of solvent* as in the solution, which will then be at a lower pressure than that of the solution by the amount P_{osm}.

Using the B_m's already discussed in section 4.8 we shall now proceed to show that

$$\ln (\zeta_a/\rho_a) = -\rho_a \sum_{m \geqslant 2} m_a B_m \boldsymbol{\rho}^m, \tag{6.9.13}$$

and

$$P_{\text{osm}} = kT[\rho - \sum_{m \geqslant 2} (m-1) B_m \boldsymbol{\rho}^m]. \tag{6.9.14}$$

However, the general rigorous derivation of these becomes quite complicated. One has always the complication of the combinatorial calculation given in (4.4) for one component, which is more involved for more components. In addition the very definition of the B_m's becomes non-trivial if we have both of the complications due to many components *and* those due to a failure of the assumption of pair additivity in the potentials of average force. For one solute component one can always define B_m by the implicit relation, for $n \leqslant 2$,

$$n^2 b_n = \left\{ \sum_{\nu_m \geqslant 0} \left[\prod_m (nB_m)^{\nu_m}/\nu_m! \right] \right\}_{\Sigma(m-1)\nu_m = n-1} \tag{6.9.15}$$

for which the explicit inverse is

$$B_m = \left\{ \sum_{\nu_n \geqslant 0} (-)^{1+\Sigma_n} \frac{[m - 2 + \Sigma \nu_n]!}{m!} \prod_{n \geqslant 2} \frac{(nb_n)^{\nu_n}}{\nu_n!} \right\}_{\underset{n}{\Sigma}(n-1)\nu_n = m-1.} \tag{6.9.16}$$

In the multicomponent case one can define the component potentials, $\omega_{\mathbf{v}}\{\mathbf{v}\}$, of average force by (6.5.2) and introduce quantities,

$$f_n\{\mathrm{n}\} = [\exp - \omega_n(\boldsymbol{n})/kT] - 1, \tag{6.9.17}$$

so that

$$\exp - (W_N\{N\}/kT) = \prod_{n \geqslant 2} \prod_{\{n\}_N} [1 + f_n\{n\}_N]. \tag{6.9.18}$$

Now develop the product. Inventing different colored n-pointed stars for the f_n's one can imagine diagrams for every product, and again the sum of all possible products corresponding to diagrams between the set $\boldsymbol{m}$ of vertices with no nodal vertex will form an irreducible function, $\boldsymbol{\Gamma}_m(\boldsymbol{m})$, eq. (4.8.4), whose integral in V over all $d\{\boldsymbol{m}\}$ divided by V and by $\boldsymbol{m}!$ will be B_m, eq. (4.8.5).

The next step would be to find the combinatorial factor for the number of ways of forming a product belonging to b_n from a given product of power ν_m of the B_m's, corresponding to $C(m_1, \ldots, m_\nu, \ldots)$ of eq. (4.4.8) in the one-component case. From there the steps would follow a course similar to the derivation in section 4.4 using the Cauchy integral method corresponding to (4.4.13). This procedure has been carried out by Fuchs for the general case.† We short circuit this lengthy and complicated argument by using a proof that appears to be completely compelling, even if somewhat cryptic.

For one component

$$B_m = m^{-1}\beta_{m-1}, \tag{6.9.19}$$

and from (4.2.12) and (4.4.21) we have the two equations,

$$P_{\text{osm}} = kT \sum_{n\geqslant 1} b_n z^n \tag{6.9.20}$$

$$= kT[\rho - \sum_{m\geqslant 2} (m-1)B_m \rho_m]. \tag{6.9.20'}$$

Now suppose there are distinguishable molecular species a, b, . . ., but the potentials of average force $\Omega_n\{n\}$ were completely independent of the species of molecules in the set of n, a case that is closely approached by molecules containing differing isotopic species of moderately heavy atoms. Let x_a be the mol-fraction of species a, $x_a = N_a/N$, so that

$$\sum_a x_a = 1, \tag{6.9.21}$$

and let z_a be the activity of species a.

The Gibbs free energy G has, for the mixture of mechanically identical molecules, the additive term $+ kT\sum_a N_a(\ln N_a/N)$ due to the entropy of mixing (see section 2.6), and with $\mu_a = (\partial G/\partial N_a)_T$ we have $\mu_a = \mu + kT \ln x_a$ where μ is the chemical potential per molecule of a one-component system of the same temperature and pressure. With (4.1.3) for z we have

$$z_a = x_a z, \tag{6.9.22}$$

where z is the activity of the hypothetical one-component system. The pressure is given by (20) with the number density ρ given by

$$\rho = [\partial(P_{\text{osm}}/kT)/\partial \ln z)_T. \tag{6.9.23}$$

Now for the mixture of different species we introduce,

$$\mathbf{b}_\mathbf{n} = \frac{\mathbf{n}!}{\mathbf{n}!} b_n, \qquad \mathbf{B}_\mathbf{m} = \frac{m!}{\mathbf{m}!} B_m, \tag{6.9.24}$$

---

† K. Fuchs, *Proc. Roy. Soc.* A **179**, 408 (1942).

and

$$\begin{aligned} P_{\text{osm}} &= kT \sum_{n\geqslant 1} b_n z^n \\ &= kT \sum_{n\geqslant 1} b_n z^n \left\{ \sum_{n_a} \frac{n!}{\prod\limits_a n_a!} \prod_a x_a{}^{n_a} \right\}_{\Sigma n_a = n} \\ &= kT \sum_{n\geqslant 1} b_n z^n (\Sigma x_a)^n \\ &= k \sum_{n\geqslant 1} b_n z^n, \end{aligned} \tag{6.9.25}$$

the same as for one component of activity z. The number density of species a is now

$$\begin{aligned} \rho_a &= z_a \left(\frac{\partial}{\partial z_a}\right)\left(\frac{P_{\text{osm}}}{kT}\right) = x_a \left(z \frac{\partial}{\partial z}\right)_T \frac{P_{\text{osm}}}{kT} \\ &= x_a \rho, \end{aligned} \tag{6.9.26}$$

which, of course, is required if x_a is to be the mol-fraction of a.

We may now use (20′) for the pressure of the one-component system. We assume that, for the multicomponent system we can expand P_{osm} as a power series in the ρ_a's, and noting that, analogously to eq. (25),

$$\begin{aligned} B_m \rho^m &= B_m \rho^m (\sum_a x_a)^m \\ &= B_m \rho^m \left\{ \sum_{m_a} \frac{m!}{\prod\limits_a m_a!} \prod_a x_a{}^{m_a} \right\}_{\Sigma m_a = m} \\ &= \{\sum_{m_a} B_{\mathbf{m}} \prod_a \rho_a{}^{m_a}\} \Sigma m_a = m \\ &= \{\sum B_{\mathbf{m}} \boldsymbol{\rho}^{\mathbf{m}}\}_{|\mathbf{m}| = m}, \end{aligned} \tag{6.9.27}$$

we derive (20′) from (20) for the case of indistinguishable forces between molecules of different species. But if this is correct in the asymptotic limit it must be the correct general equation, provided an expansion with coefficients of the nature of the $B_{\mathbf{m}}$'s exists.

6.10. Moments and thermodynamic quantities

There is a bewildering variety of relations between thermodynamic quantities such as the first, second, and higher derivatives of P/kT with respect to $1/kT$, $\ln z_a$, $\ln z_b$, etc., and integrals of functions involving the probability densities, $\rho_n\{\mathbf{n}\}$. We shall demonstrate and discuss a few of these.

The probability density, $W_N\{N\}$, in the coordinate space, $\{N\}$, of finding exactly the set N of molecules in volume V, and of finding these at the position $\{N\}$, is

$$W_N\{N\} = \mathbf{z}^N \exp - \frac{1}{kT}(PV + U_N\{N\}), \tag{6.10.1}$$

normalized so that

$$\sum_{N \geqslant 0} \iint_V \cdots \int \frac{1}{N!} W_N\{N\} d\{N\} = 1. \tag{6.10.2}$$

The normalization (2) must hold for all values of the parameters $(1/kT)$, $\ln z_a$, $\ln z_b$, that determine the state so that, for instance,

$$\begin{aligned} &\sum_{N>0} \iint_V \cdots \int \frac{1}{N!} (\partial W_N / \partial \ln z_a)_{T,z_b} d\{N\}, \\ &= \sum_{N>0} \iint_V \cdots \int \frac{1}{N!} W_N \left\{ -V \left[\frac{\partial(P/kT)}{\partial \ln z_a} \right]_{T,z_b} + N_a \right\} \\ &= -V[\partial(P/kT)/\partial \ln z_a]_{T,z_b} + \bar{N}_a = 0, \end{aligned} \tag{6.10.3}$$

with $\bar{N}_a$ the average number of molecules of species a in V. Taking the derivative with respect to $(1/kT)$ gives

$$-V[\partial(P/kT)/\partial(1/kT)]_z - \bar{U} = 0, \tag{6.10.4}$$

with $\bar{U}$ the average potential energy in V. Define

$$\bar{u} = \bar{U}/\bar{N} \tag{6.10.5}$$

as the average potential energy per molecule. Since $\bar{u} = 0$ in the perfect gas and the kinetic energy is classically independent of density, we have

$$u(z,T) = \bar{N}^{-1}[E(z,T) - E_{\text{perfect gas}}] \tag{6.10.6}$$

for a classical system. With (5) we have, from (3) and (4),

$$[\partial(P/kT)/\partial \ln z_a]_{T,z_b} = x_a \rho, \tag{6.10.7}$$

$$T[\partial(P/kT)/\partial T]_z = \rho(\bar{u}/kT), \tag{6.10.8}$$

where eq. (6.3.15) in a σ-component system gives

$$\bar{u} = \frac{1}{2} \sum_{a=1}^{a=\sigma} \sum_{b=1}^{b=\sigma} x_a x_b \bar{u}_{ab}. \tag{6.10.9}$$

Equation (7) above has been used repeatedly as a thermodynamic equation, and, since $(d \ln z)_{T,z_b} = (1/kT)(d\mu_a)_{T,z_b}$ the proof given here is essentially the same as that given in section 1.3, eq. (1.3.12). The relation (8) is new, and can be derived straightforwardly but tediously from (7) and (1.3.19) that $(\partial P/\partial T)_\mu = S/V$.

Using (7) and (8) for derivatives of P/kT we can write,

$$(\partial W_N/\partial \ln z_a)_{T,z_b} = W_N[N_a - \bar{N}_a(z,T)], \tag{6.10.10}$$

$$[\partial W_N/\partial(1/kT)]_z = - W_N[U_N\{N\} - \bar{U}(z,T)], \tag{6.10.11}$$

in which the averages, $\bar{N}_a$, $\bar{U}$, depend on z and T, but the two functions, N_a, and $U_N\{N\}$ do not. The quantity N_a in (10) is a trivial function

$$N_a = \sum_{ia=1a}^{ia=N_a} 1, \tag{6.10.12}$$

namely the sum over all molecules of species a of unity, but it can be, and will be, regarded as a function in what follows.

In general if $\Psi_N\{N\}$ is any function of the coordinates of the set N of molecules, we can define its average value $\bar{\psi}$, per molecule, as

$$\bar{\psi} = \bar{N}^{-1}\Psi = N^{-1} \sum_{N \geqslant 0} \frac{1}{N!} \int\!\!\int_V \cdots \int W_N \Psi_N d\{N\}. \tag{6.10.13}$$

For two functions, Ψ_N, χ_N we define product moments by

$$\begin{aligned} I_{\Psi,\chi} &= I_{\chi,\Psi} \\ &= \bar{N}^{-1} \sum_{N \geqslant 0} \frac{1}{N!} \int\!\!\int_V \cdots \int W_N[\Psi_N - \bar{\Psi}][\chi_N - \bar{\chi}] d\{N\} \\ &= N^{-1}[<\Psi\chi> = \bar{\Psi}\bar{\chi}], \end{aligned} \tag{6.10.14}$$

where $<\Psi\chi>$ is the average value of the product function, $\Psi_N\chi_N$.

Since, necessarily,

$$\sum_{N>0} \frac{1}{N!} \int\!\!\int_V \cdots \int W_N[\Psi_N - \bar{\Psi}] d\{N\} = 0, \tag{6.10.15}$$

it follows that if x is any paramenter on which W_N or Ψ_N or both depend we can, by differentiation under the integral obtain,

$$\begin{aligned} &(\partial\bar{\Psi}/\partial x) - <\partial\Psi_N/\partial x> \\ &\quad = \sum_{N>0} \frac{1}{N!} \int\!\!\int_V \cdots \int W_N[\Psi_N - \bar{\Psi}][\partial \ln W_N/\partial x] d\{N\} \\ &\quad = \bar{N} I_{\Psi,\ \partial \ln W/\partial x} \end{aligned} \tag{6.10.16}$$

we shall let x be $\ln z_b$ and $1/kT$ consecutively, using (10) and (11) for $\partial \ln W_N/\partial x$. For Ψ_N choose three different functions. The first is N_a with $\bar{N}_a(z,T) = V[\partial(P/kT)/\ln z_a]_{z_b,T}$. The second is U_N with $\bar{U}(z,T) = -[\partial(P/kT)/\partial(1/kT)]_z$. The third is the function U_N', defined by eq. (6.6.10) and given, for the simple case of a sum of pair mutual potentials by (6.6.14),

$$U_N' = \sum_{N>i>j>1} \sum r_{ij}[du(r_{ij})/dr_{ij}]. \tag{6.10.17}$$

From (6.6.9) we have

$$\begin{aligned} \bar{U}' &= 3\bar{N}kT - 3PV \\ &= 3VkT[\rho - (P/kT)], \end{aligned} \tag{6.10.18}$$

$$(\partial\bar{U}'/\partial \ln z_a)_{T,z_b} = 3VkT(\partial\rho/\partial \ln z_a)_{z_b,T} - 3\bar{N}_a kT, \tag{6.10.19}$$

$$[\partial\bar{U}'/\partial(1/kT)]_z = 3VkT\left\{P - \rho kT\left(1 + \frac{\bar{u}}{kT}\right) + \left[\frac{\partial\rho}{\partial(1/kT)}\right]_z\right\}. \tag{6.10.20}$$

None of the functions, N_a, U_N, or U_N' depend, as functions of $\{N\}$, on $\ln z_b$ or $1/kT$. We find, then, from (16)

$$\frac{1}{\rho}\left[\frac{\partial^2(P/kT)}{\partial \ln z_a\, \partial \ln z_b}\right]_T = \frac{1}{\rho}\left(\frac{\partial\rho_a}{\partial \ln z_b}\right)_{z_a,T} = \frac{1}{\rho}\left(\frac{\partial\rho_b}{\partial \ln z_a}\right)_{z_b,T} = I_{N_a,N_b}, \tag{6.10.21}$$

$$\frac{T}{\rho}\left[\frac{\partial^2(P/kT)}{\partial \ln z_a\, \partial T}\right] = \frac{1}{\rho kT}\left[\frac{\partial(\rho\bar{u})}{\partial \ln z_b}\right]_{z_a,T} = \frac{T}{\rho}\left(\frac{\partial\rho_a}{\partial T}\right)_z = \frac{1}{kT} I_{N_a,U}, \tag{6.10.22}$$

$$\frac{1}{\rho}\left(\frac{1}{kT}\right)^2\left\{\frac{\partial^2(P/kT)}{[\partial(1/kT)]^2}\right\}_z = \frac{1}{kT}\,\frac{T}{\rho}\left[\frac{\partial(\rho\bar{u})}{\partial T}\right]_z = \left(\frac{1}{kT}\right) I_{U,U}, \tag{6.10.23}$$

$$3\left[\left(\frac{\partial \ln \rho}{\partial \ln z_a}\right)_{z_b,T} - x_a\right] = 3\,[I_{N,N_a} - x_a] = \frac{1}{kT} I_{N_a,U'}, \tag{6.10.24}$$

$$\begin{aligned} &3\left[\frac{P}{\rho kT} - 1 - \frac{\bar{u}}{kT} - \frac{T}{\rho}\left(\frac{\partial\rho}{\partial T}\right)_z\right] \\ &\quad = 3\left[\frac{P}{\rho kT} - 1 - \frac{\bar{u}}{kT} - \frac{1}{kT} I_{N,U}\right] = \left(\frac{1}{kT}\right)^2 I_{U',U}. \end{aligned} \tag{6.10.25}$$

By using the coordinate transformation to dimensionless coordinates, discussed in section 6.6, and the operator $V(\partial/\partial V)_{T,z}$ we can also find expressions involving $I_{U',U'}$.

We now examine in detail the evaluation of the averages and product moments. We have the trivial function N_a of (12) for which the products are

$$[N_a\{N\}]^2 = \sum_{ia=1a}^{i=N} 1 + 2 \sum_{Na \geqslant ia > ja \geqslant 1a} \sum 1, \tag{6.10.26}$$

$$[N_a\{N\}][N_b\{N\}] = \sum_{ia=1a}^{ia=N_a} \sum_{jb=1b}^{jb=Nb} 1, \qquad (a \neq b), \tag{6.10.26$'$}$$

and the sums of pairs, U_N and U_N',

$$\Psi_N\{N\} = \sum_{N \geqslant i > j \geqslant 1} \sum \psi_{ij}(r_{ij}) \tag{6.10.27}$$

in which the function, ψ_{ij}, depends on the species to which i and j belong. The product $N_a\Psi_N$ is

$$\begin{aligned} N_a\{N\}\Psi_N\{N\} = 2 &\sum_{N_a \geqslant i > j \geqslant 1a} \sum \psi_{aa}(r_{ij}) + \sum_b{}' \sum_{ia=1a}^{ia=N_a} \sum_{jb=1b}^{jb=N_b} \psi_{ab}(r_{ia,jb}) \\ &+ \sum_{N > i > j > k \geqslant 1} \sum \sum (\delta(i-a)\psi_{j,k}(r_{jk}) + \delta(j-a)\psi_{ki}(r_{ki}) \\ &+ \delta(k-a)\psi_{ij}(r_{ij})], \end{aligned} \tag{6.10.28}$$

where, by $\delta(i-a)$ we mean that $\delta(i-a)$ is unity if i is of type a and otherwise zero. The product of two pair functions is

$$\begin{aligned} [\Psi_N\{N\}][\chi_N\{N\}] = &\sum_{N > i > j > 1} \sum \Psi_{ij}(r_{ij})\chi_{ij}(r_{ij}) \\ &+ \sum_{N > i > j > k > 1} \sum \sum [\psi_{ki}\chi_{ij} + \chi_{ki}\psi_{ij} + \psi_{ij}\chi_{jk} \\ &+ \chi_{ij}\psi_{jk} + \psi_{jk}\chi_{ki} + \chi_{jk}\psi_{ki}] \\ &+ \sum_{N > i > j > k > l > 1} \sum \sum \sum [\psi_{ij}\chi_{kl} + \chi_{ij}\psi_{kl} + \psi_{ik}\chi_{jl} \\ &+ \chi_{ik}\psi_{jl} + \psi_{il}\chi_{jk} + \chi_{il}\psi_{jk}]. \end{aligned} \tag{6.10.29}$$

With the definition,

$$F_2(r)\psi_2(r) = \sum_{a=1}^{a=\sigma} \sum_{b=1}^{b=\sigma} x_a x_b F_{ab}(r)\psi_{ab}(r), \tag{6.10.30}$$

we have, for the averages,

$$<N_a> = \bar{N} x_a \rho = \int_V \rho_a dr, \tag{6.10.31}$$

$$\begin{aligned} <\Psi> &= \tfrac{1}{2} \bar{N}\bar{\psi} = \tfrac{1}{2} \bar{N}\rho \int 4\pi r^2 F_2(r)\psi(r)dr \\ &= \tfrac{1}{2} \iint_V \rho^2 F_2(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)\psi(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)d\boldsymbol{r}_1 d\boldsymbol{r}_2. \end{aligned} \tag{6.10.32}$$

For the products of averages we have terms proportional to $\bar{N}^2$ which can be written as multiple integrals,

$$<N_a><N_b> = \iint_V \rho_a \rho_b d\mathbf{r}_1 d\mathbf{r}_2, \tag{6.10.33}$$

$$<N_a><\Psi> = \iiint_V \rho^3 x_a F_2(r_{12}) \psi(r_{12}) d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3, \tag{6.10.34}$$

$$<\Psi><\chi> = \iiiint_V \rho^4 F_2(r_{12}) F_2(r_{34}) \psi(r_{12}) \chi(r_{34}) \\ d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4. \tag{6.10.35}$$

In the products of functions, eqs. (26) to (29) the sums over the largest subsets (pairs for N_a^2 and $N_a N_b$, triples for $N_a \Psi$, and quadruples for $\Psi\chi$) when multiplied by W_N, integrated over $d\{N\}$, and summed over N lead to terms of order $\bar{N}^2$. We subtract from these the appropriate product of averages, eqs. (32) to (35), *before* integration over the last subset of molecules.

The results can be expressed in terms of integrals,

$$J_{ab} = \bar{N}^{-1} \iint [\rho_{ab}(r_{12}) - \rho_a \rho_b] d\mathbf{r}_1 d\mathbf{r}_2 \\ = x_a x_b \rho \int 4\pi r^2 [F_{ab}(r) - 1] dr, \tag{6.10.36}$$

$$K_{a,\psi} = \rho^2 \iint_V \sum_{b=1}^{b=\sigma} \sum_{c=1}^{c=\sigma} x_b x_c [F_{abc}(r_b, r_c, r_{bc}) - F_{bc}(r_{bc})] \\ \psi_{bc}(r_{bc}) d\mathbf{r}_b d\mathbf{r}_c, \tag{6.10.37}$$

$$L_{\Psi,\chi} = \rho \int 4\pi r^2 [\sum_{a=1}^{a=\sigma} \sum_{b=1}^{b=\sigma} x_a x_b F_{ab}(r) \psi_{ab}(r) \chi_{ab}(r)] dr, \tag{6.10.38}$$

$$M_{\Psi,\chi} = \rho^2 \iint \sum_{a=1}^{a=\sigma} \sum_{b=1}^{b=\sigma} \sum_{c=1}^{c=\sigma} x_a x_b x_c F_{abc}(r_a, r_c, r_{ac}) \chi_{bc}(r_c) \psi_{ab}(r_a) \\ d\mathbf{r}_a d\mathbf{r}_c, \tag{6.10.39}$$

$$N_{\Psi,\chi} = \rho^3 \iiint \sum_{a=1}^{a=\sigma} \sum_{b=1}^{b=\sigma} \sum_{c=1}^{c=\sigma} \sum_{d=1}^{d=\sigma} x_a, x_b, x_c, x_d \\ \times [F_{abcd}(r_{ab}, r_{ac}, r_{ad}, r_{bc}, r_{bd}, r_{cd}) - F_{ab}(r_{ab}) F_{cd}(r_{cd})] \\ \times \psi_{ab}(r_{ab}) \chi_{cd}(r_{cd}) d\mathbf{r}_b d\mathbf{r}_c d\mathbf{r}_d, \tag{6.10.40}$$

all of which will be independent of the volume of the system if the functions $\psi(r)$, $\chi(r)$ go sufficiently rapidly to zero. One then finds,

$$I_{N_a,Nb} = \delta(a-b)x_a + x_a x_b J_{ab}, \tag{6.10.41}$$

$$I_{N,N_a} = x_a(1 + J_{a\sigma}) = x_a[1 + \sum_{b=1}^{b=\sigma} x_b J_{ab}), \tag{6.10.42}$$

$$I_{N,N} = 1 + J = 1 + \sum_{a=1}^{a=\sigma} \sum_{b=1}^{b=\sigma} x_a x_b J_{ab}, \tag{6.10.43}$$

$$I_{N_a,\Psi} = x_a(\sum_{b=1}^{b=\sigma} x_b \bar{\psi}_{ab} + \tfrac{1}{2} K_{a,\Psi}), \tag{6.10.44}$$

$$I_{N,\Psi} = \bar{\psi} + \tfrac{1}{2} K_{\sigma,\Psi} = \bar{\psi} + \tfrac{1}{2} \sum_{a=1}^{a=\sigma} x_a K_{a,\Psi}, \tag{6.10.45}$$

$$I_{\Psi,\chi} = \tfrac{1}{2} L_{\Psi,\chi} + M_{\Psi,\chi} + \tfrac{1}{4} N_{\Psi,\chi}. \tag{6.10.46}$$

Setting Ψ_N, χ_N equal to U_N or U_N' we find, through eqs. (21) to (25) relations between thermodynamic quantities and the various integrals of eqs. (36) to (40). Only few examples will be discussed, and these for a one-component system.

The isothermal compressibility, κ, is

$$\kappa = \rho^{-1}(\partial \rho / \partial P)_T, \tag{6.10.47}$$

and, for one component, using (21), with (7),

$$\begin{aligned} \rho^{-1}(\partial \rho / \partial \ln z)_T &= \rho^{-1}(\partial \rho / \partial P)_T \, kT \, [\partial (P/kT) / \partial \ln z]_T \\ &= \rho k T \kappa = I_{N,N} = 1 + J, \end{aligned} \tag{6.10.48}$$

or with (36) for J,

$$\rho \int 4\pi r^2 [F(r) - 1] dr = \rho k T \kappa - 1. \tag{6.10.49}$$

In the limit of the perfect gas, $\rho \to 0$ we find $\rho k T \kappa = 1$, $\kappa = (1/\rho k T)$, the correct result. For very highly compressed material the compressibility κ is very small compared with $1/\rho k T$ and the integral on the left of (49) is close to -1. If $F(r)$ were given by $F(r) = 0$, $r \leqslant r_0$, $F(r) = 1$, $r > r_0$ we would then have $(4\pi/3) r_0^3 \cong 1/\rho$, namely around each molecule there is an excluded volume equal to the volume per molecule in the system, and beyond this volume the probability density has the normal constant value ρ. Actually this is not the form of $F(r)$ at high compression. However, the fact that κ must be positive, $I_{NN} \geqslant 0$, requires that for all systems

$$\rho \int 4\sigma r^2 [F(r) - 1] dr \geqslant -1. \tag{6.10.50}$$

The limit $\kappa \to 0$ is found only with $I_{NN} = \bar{N}^{-1}[\langle N^2 \rangle - \bar{N}^2] = 0$, and this can only be the case if the number of molecules in V is precisely determined, and fluctuations are zero. This condition, of course, is most closely approached at highly compressed densities of packing.

The dimensionless quantity

$$\begin{aligned} c_{zu} &= k^{-1}[\partial \bar{u}/\partial T]_z \\ &= (1/kT)(T/\rho)[\partial(\rho\bar{u})/\partial T]_z - (\bar{u}/kT)(T/\rho)(\partial\rho/\partial T)_z \end{aligned} \quad (6.10.51)$$

is, from (22) and (23),

$$c_{z,u} = \left(\frac{1}{kT}\right)^2 [I_{U,U} - \bar{u} I_{N,U}]. \quad (6.10.52)$$

This is, in units of Boltzmann's k, the heat capacity per molecule due to potential energy alone *at constant activity*, a rather horrible thermodynamic quantity. The similar dimensionless heat capacity per molecule at constant volume,

$$c_V = (kN)^{-1}(\partial E/\partial T)_{N,V}, \quad (6.10.53)$$

has a classical contribution of 3/2 due to kinetic energy, and a contribution, $c_{V,u}$, due to potential energy,

$$c_{V,u} = c_V - \tfrac{3}{2}kT. \quad (6.10.54)$$

By tedious but straightforward manipulation one finds that

$$c_{V,u} = c_{z,u} + \frac{\bar{u}}{kT}\left[\frac{T}{\rho}\frac{\partial T}{\partial \rho}_z\right] - \frac{1}{\rho kT\kappa}\left[\frac{T}{\rho}\frac{\partial \rho}{\partial T}_z\right]^2 \quad (6.10.55)$$

from which, with (21), (22) and (52), one finds

$$c_{V,u} = \frac{1}{kT}{}^2[I_{U,U} - (I_{N,U})^2 (I_{N,N})^{-1}]. \quad (6.10.56)$$

6.11. Open and closed system moments

The manipulations of the last section were made on the equations for an "open" system, for which the ensemble is defined by the variables z and T. Were the same manipulations carried out on a "closed" system for which the ensemble is one of fixed number set N and T, one would have come to the result that $c_{V,u} = I_{U,U}/(kT)^2$ instead of (6.10.56). Furthermore $I_{U,U}$ would have been given in terms of the same integrals, $L_{U,U}$, $M_{U,U}$, and $N_{U,U}$ by (6.10.46) and (6.10.38) to (6.10.40)! The answer to this apparent discrepancy lies in a subtle difference in the approach of $\rho_{n+m}(\{n\} + \{m\})$ to $\rho_n\{n\}\rho_m\{m\}$ when the distances between the subsets $\{n\}$ and $\{m\}$ become large.

The difference in the predictions is seen most clearly and simply in the expression for the second moment $I_{N,N}$, and this case only will be discussed in detail. The probability density $\rho_2(\boldsymbol{r}_1,\boldsymbol{r}_2)$ is that of finding simultaneously one molecule at r_1 and another at r_2. If this is integrated over $d\boldsymbol{r}_1 d\boldsymbol{r}_2$ in V it gives twice the average number of pairs since molecule i may be at $\boldsymbol{r}_1$, j at $\boldsymbol{r}_2$ *or* i at $\boldsymbol{r}_2$ and j at $\boldsymbol{r}_1$. In a fluid for which $\rho_2(\boldsymbol{r}_1,\boldsymbol{r}_2)$ depends only on $r_{12} = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$ one then has

$$\iint_V \rho_2(r_{12}) d\boldsymbol{r}_1 d\boldsymbol{r}_2 = <N(N-1)> = <N^2> - \bar{N} \qquad (6.11.1)$$

and, since $\bar{N} = \int \rho dr$,

$$\begin{aligned} &\bar{N}^{-1} \iint [\rho_2(r_{12}) - \rho^2] dr_1 dr_2 \\ &\quad = \rho \int 4\pi r^2 [F_2(r) - 1] dr = \bar{N}^{-1}[<N^2> - \bar{N}^2 - \bar{N}] \\ &\quad = I_{NN} - 1, \end{aligned} \qquad (6.11.2)$$

which is a quick and easy derivation of (6.10.43), that $I_{NN} = 1 + J$. Now in the ensemble of closed systems the number N of molecules is specified for each system, the "average" number, $\bar{N} \equiv N$ is exact and the average of N^2 is exactly $N^2 = \bar{N}^2$. In short $I_{N,N} = N^{-1}[N^2 - N^2] = 0$ in the closed system ensemble. We then have, for the closed system

$$\rho \int 4\pi r^2 \{[F(r)]_{\text{closed}} - 1\} dr = J_{\text{closed}} = -1, \qquad (6.11.3)$$

instead of (6.10.49) that $J = \rho k T \kappa - 1$.

The distinction is not difficult to understand. In the closed system the number N of molecules in V is fixed, and the density is $\rho = N/V$. Now we hold one molecule fixed at $\boldsymbol{r}_1$ and look for the probability density of finding a second at $\boldsymbol{r}_2$ with $|\boldsymbol{r}_1 - \boldsymbol{r}_2|$ so large that there is no *direct* effect of the one at r_1. This probability density will be that of finding one of the $N - 1$ *other* molecules in some volume $V - b$, where b is a volume of the order of that of a molecule, namely that excluded from occupancy by another molecule due to the first at $\boldsymbol{r}_1$. We then have

$$\begin{aligned} \lim_{r_{12} >> r_0} \left[\rho_2(r_{12}) = \rho \frac{N-1}{V-b} \right] \\ = \rho^2 \left[1 - \frac{1}{N} \frac{1 - b\rho}{1 - N^{-1} b\rho} \right] \\ = \rho^2 [1 - N^{-1}(1 - b\rho) + \text{order } N^{-2}]. \end{aligned} \qquad (6.11.4)$$

The integral of $[\rho_2(r_{12}) - \rho^2] d\boldsymbol{r}_2$ over the volume V includes a short range term where $F_2(r)$ differs considerably from unity, given by the J

of the open ensemble plus a term due to the integral over V of the small term $-\rho^2 N^{-1}(1 - b\rho)$. The integral of this over V is $\rho(1 - b\rho)$. The next integration of $d\boldsymbol{r}_2$ over V and division by N removes the factor ρ, so that, from (3), $J_{\text{closed}} = J - (1 - b\rho) = -1$, and with (6.10.48), $J = \rho kT\kappa - 1$, we have that

$$\rho kT\kappa = 1 - b\rho. \tag{6.11.5}$$

Again in the perfect gas case, $\rho \to 0$ we have $b\rho << 1$ and $\rho kT\kappa = 1$, whereas for the other extreme of very high compression $b\rho \cong 1$ and $\kappa \to 0$.

The other integrals, K and N of eqs. (6.10.37) and (6.10.40) have similar limits of the order of N^{-1} instead of zero in the case of the closed systems.

The open systems, on the other hand, are defined as a finite volume V with walls permeable to molecules in equilibrium with an infinite reservoir of the same temperature and activity. The value of $\rho_2(r_{12})$ is now given by (4) in the limit that the N refers to the infinite reservoir, namely $N \to \infty$, whereas the volume over which the integrals for $I_{N,N}$, etc., are extended remains finite. Expressed less formally, if we fix the position of a molecule at $\boldsymbol{r}_1$ then, at a sufficient distance the probability density of finding a second molecule at r_2 is determined by the density of the infinite reservoir, and is ρ.

6.12. Integral equations

The series developments in terms of cluster and irreducible integrals in powers of the differences of activities all diverge at phase transitions. They have never been used effectively to go from the true mutual potentials between molecules, in which one starts at zero activity, to obtain information about liquids below the critical temperature. The only available method appears to be that of the use of some one or other of a considerable number of integral equations.

Formally the various integral equations are all derivable from the set of equations exemplified by (6.7.8) giving $G_n = (\rho/z)^n[\exp(VP/kT)]F_n$ at activity set $\boldsymbol{z}$ in terms of integrals over $d\{N\}$ of the functions G_{n+N} for some other activity set $\boldsymbol{y}$. Now from (6.7.1) for G_n with (6.4.1) for $\ln F_n$ and (6.5.1) one has for $\ln G_N$,

$$\ln G_N = (VP/kT) + \sum_{i=1}^{i=N} (\ln \rho i - \ln z_i)$$
$$- \sum_{N \geqslant i > j \geqslant 1} \sum \omega_{ij}(r_{ij})/kT - \sum_{n \geqslant i > j > k \geqslant 1} \sum \sum \omega_{ijk}/kT - \ldots, \tag{6.12.1}$$

namely a constant plus sums of terms that depend on small subsets of molecules, and which, with the Kirkwood superposition assumption may be taken to stop with the pair functions ω_{ij}/kT.

We write (6.7.8) as

$$G_N(\mathbf{z},\{\mathbf{n}\}) = \sum_{N\geqslant 0} \frac{(\mathbf{z}-\mathbf{y})^N}{N!} \iint \cdots \int_V G_{n+N}(\mathbf{y},\{\{\mathbf{n}\}+\{N\}\})d\{N\}, \tag{6.12.2}$$

and operate on both sides by some type of differential operator which acts on the functions ω_n/kT. We will have, on the left, $G_n(z)$ times a sum of terms involving functions, $\psi_\nu(z\{\nu\}_n)$ of small subsets $\{\nu\}_n$ of ν molecules of the set $\mathbf{n}$. These functions arise from operation on the functions $[\omega_\nu(z)/kT]$ at activity set $\mathbf{z}$. On the right under the integral one has a similar sum of functions, $\psi_{\nu+m}(\mathbf{y},\{\{\nu\}_n+\{m\}_N\})$, of subset $\boldsymbol{\nu}+\mathbf{m}$ from both $\mathbf{n}$ and N, obtained by operation on $[\omega_{\nu+m}(\mathbf{y}/kT]$.

The essential feature is to assume that no non-zero terms exist for large $\{\mathbf{m}\}_N$, and preferably that the differential operator can be so chosen that m never exceeds unity. For each term, $\psi_{\nu+m}$, integrate over all the molecules of N except those in $\{\mathbf{m}\}_N$. Multiply by the number of such terms in N, namely by N_a if $\{\mathbf{m}\}_N$ is only one molecule of species a, or in general by $N!/(N-\mathbf{m})!\mathbf{m}!$. From (6.7.8) one has on the right sums of integrals over small sets $\mathbf{m}$ of molecules,

$$\iint \cdots \int \frac{1}{m!} G_{n,m}(\mathbf{z},\{\{n\}+\{\mathbf{m}\}\})\psi_{\nu+m}(\mathbf{y},\{\{\nu\}_n+\{\mathbf{m}\}\})d\{\mathbf{m}\}, \tag{A}$$

in which the G-function is now that of activity $\mathbf{z}$ whereas the ψ-function is obtained by operation on $[\omega_{\nu+m}(\mathbf{y})/kT)$. There will also be terms for which $m \equiv 0$, on the right, namely,

$$G_n(\mathbf{z}\{\mathbf{n}\})\psi_\nu(\mathbf{y},\{\nu\}_n). \tag{B}$$

Divide both sides of the result by $G_n(\mathbf{z})$, define

$$Q_{n,m}(z,\{\mathbf{n}\},\{\mathbf{m}\}) = [G_n(\mathbf{z},\{\mathbf{n}\})]^{-1}[G_{n+m}(\mathbf{z}\{\mathbf{n}+\mathbf{m}\})], \tag{6.12.3}$$

and carry the terms (B) to the left-hand side. One finds

$$(\Sigma\{\nu\}_n[\psi_\nu(\mathbf{z},\{\nu\}_n) - \psi_\nu(\mathbf{y},\{\nu\}_n)]$$
$$= (\Sigma\{\nu\}_n \sum_{M\geqslant 1} \frac{1}{\mathbf{m}!} \iint \cdots \int Q_{n,m}(\mathbf{z}\{\mathbf{n}\},\{\mathbf{m}\})\psi_{\nu+m}(\mathbf{y},\{\{\nu\}_n+\{\mathbf{m}\})d\{\mathbf{m}\}. \tag{6.12.4}$$

From here various procedures can be used leading to a wide variety of utterly different appearing equations.

Probably the most appealing of these is that which leads to the Yvon Born–Green equation, since this can also be written solely from the knowledge, proved in section 6.4, that $\Omega_n(z,\{\mathbf{n}\})$ is the potential of average force. We assume a one-component system and choose $n = 2$ so that for a fluid with $\omega_1 = 0$, $\Omega_2 = \omega_2(r_{ij})$. The average force, $f_{xi}(\mathbf{r}_i,\mathbf{r}_j) = -(\partial/\partial x_i)\omega_2(|\mathbf{r}_i - \mathbf{r}_j|)$ is that acting along the x-axis on a molecule at $\mathbf{r}_i$ when another molecule is at $\mathbf{r}_j$. This force will be the direct force, $-(\partial/\partial x_i)u(|\mathbf{r}_i - \mathbf{r}_j|)$ due to molecule j plus an indirect term due to the other molecules whose positions are influenced by the presence of molecules at $\mathbf{r}_i$ and $\mathbf{r}_j$. Assume the potential is strictly a sum of pair potentials, so that the indirect force is $-(\partial/\partial x_i)u(\mathbf{r}_i,\mathbf{r}_k)$ due to any molecule at r_k multiplied by the probability density, $\rho_3(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k)/\rho_2(\mathbf{r}_i,\mathbf{r}_j)$ that a molecule be at $\mathbf{r}_k$, and integrated over all $\mathbf{r}_k$. We can therefore write,

$$\begin{aligned}\partial\omega_2(|\mathbf{r}_i - \mathbf{r}_j|)/\partial x_i - \partial u(|\mathbf{r}_i - \mathbf{r}_j|)/\partial x_i \\ = \int \frac{\rho_3(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k)}{\rho_2(\mathbf{r}_i,\mathbf{r}_j)} \frac{\partial u(|\mathbf{r}_i - \mathbf{r}_k|)}{\partial x_i} d\mathbf{r}_k.\end{aligned} \tag{6.12.5}$$

Now formally (5) follows from (4) if we choose a one-component system, set $y = 0$, $n = 2$, and, with any z, the original differential operator to be $-kT\,\partial/\partial x_i$ operating on functions of the coordinate of molecule i. With the potential a sum of pair functions the only term on the right of (4) is that with $m = 1$, $\psi_{i+k} = (\partial/\partial x_i)u(|\mathbf{r}_i - \mathbf{r}_k|)$, and $Q(z,\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k) = \rho_3(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k)[\rho_2(\mathbf{r}_i,\mathbf{r}_j)]^{-1}$.

In (5) use the Kirkwood closure, $F_3(r_1,r_2,r_3) = F_2(r_{12})F_2(r_{23})F_2(r_{31})$, with $F_2 = \exp(-\omega_2/kT)$. Since

$$(\partial/\partial x_i)u(|\mathbf{r}_i - \mathbf{r}_j|) = [(x_i - x_j)/r_{ij}]du(r_{ij})/dr_{ij} \tag{6.12.6}$$

one has an integral differential equation,

$$\begin{aligned}(x_i - x_j)[d\omega_2(r_{ij})/dr_{ij}] - [du(r_{ij}/dr_{ij}] \\ = \rho \int \left\{\exp - \frac{1}{kT}[\omega_2(r_{ik}) + \omega_2(r_{jk})]\right\} \frac{x_i - x_k}{r_{ik}} \frac{du(r_{ik})}{dr_{ik}} dr_k,\end{aligned} \tag{6.12.7}$$

involving the "known" function $u(r)$ and one unknown function $\omega_2(r)$. Without the Kirkwood assumption there would be an additional unknown term $-(1/kT)\omega_3(r_{ij},r_{jk},r_{ki})$ in the exponent. One can of course attempt consecutive approximations solving (7) as a first approximation and then using (4) with $n = 3$ assuming $\omega_4 = 0$ to solve for ω_3. No procedures of this nature appear to have been successfully carried through.

We shall not discuss in detail any of the many other manipulations that can be performed on (4) to lead to various approximate equations which have been proposed or used. We can describe a few qualitatively.

Firstly using the same operator of $-kT(\partial/\partial x_i)$ on a molecule i of the set $\boldsymbol{n}$, one can write (4) with $\boldsymbol{z} = 0$ and y some activity set of interest. In this case the expression under the integral has $Q_{n,m}$ at $z = 0$ which now contain only the "known" pair potentials $u(r)$. The unknown function $d\omega_2(r)/dr$ then occurs under the integral on the right. If again Kirkwood superposition is used the sum in (4) stops with $m = 1$. Unlike the case of equation (7)—in which the approximation is made in the function under the integral with $m = 1$, but the sum strictly terminates at $m = 1$ if U_N is a sum of pair terms—we now have the correct function for $m = 1$, but the approximation consists in terminating the sum. The fact that the unknown appears linearly in the integral is an advantage, but numerical attempts at solution appear to lead to poorer results than (7).

Kirkwood† derived and used an equation that can be obtained readily from (4) as follows. Choose a two-component system of molecules of type a and b. Let $y_a = y_b = 0$, and choose z_a to be that of a dense fluid, say the liquid, but keep z_b vanishingly small. Let $u_{aa}(r) = u(r)$ be the pair potential for which one seeks the corresponding pure liquid pair potential of average force $\omega_2(r) = \omega_{aa}(4)$ at $z = z_a$. Now set $u_{ab}(r) = \lambda u(r)$ so that at $\lambda = 1$ the molecules of type b are identical to those of type a, whereas at $\lambda = 0$ they interact with nothing and must have $\omega_{ab}(r) = 0$. Now choose the differential operator to be $\partial/\partial\lambda$ and take $n = 2$, $\boldsymbol{n} = a,b$ in (4). Since $z_b \rightarrow 0$ the terms $Q_{ba,b}$ are vanishingly small, and the sum over m involves only the single term $m = a$, with the Kirkwood approximation that

$$Q_{ba,a} = \rho \exp - \frac{1}{kT} [\omega_{ab}(\lambda, r_{ik}) + \omega_{aa}(r_{jk})]. \tag{A}$$

The $\psi_{ab}(y = 0, r)$ is now $u(r)$ rather than the force, as in the term for y on the left of (4), but $\psi_{ab}(z,r)$ on the left is $\partial\omega_{ab}(\lambda, r_{ik})/\partial\lambda$. The solutions have considerable resemblance to the experimentally determined functions.

---

† J. G. Kirkwood, *J. Chem. Phys.* **3**, 300 (1935).

Other equations that appear utterly different in form can be derived from (4) as follows. Write eq. (4) twice, once with $y = 0$, $z = \zeta$, and once with $y = \zeta$, $z = 0$. The expressions on the left are the same in the two cases, except for a difference in sign, so that the sum of the two different right-hand sums of integrals must equal zero. Now eliminate one of the two sets $\psi_{\boldsymbol{\nu}}$, say that at ζ, from under the integral by using its equation obtained from (4) involving $Q_{n,m}(\zeta)$ and $\psi_{\boldsymbol{\nu}+\boldsymbol{m}}(y = 0)$. One then has sums of integrals of products of $Q_{n,m}(y = 0)\ Q_{\boldsymbol{\nu},k}(\zeta)\psi_{u,k}(y = 0)$ set equal to zero. Now this terrible appearing sum of integrals must equate to zero whatever differential operator is used.

Suppose the original potential energy were

$$U_N\{N\} = U_N{}^{(0)} + \lambda \Delta U_N, \tag{6.12.8}$$

with

$$\Delta U_N = (\sum_{\nu \geqslant 0} \{\boldsymbol{\nu}\})_N \psi_{\boldsymbol{\nu}}\{\boldsymbol{\nu}\}_N \tag{6.12.9}$$

and we chose the differential operator to be $\partial/\partial\lambda$ at $\lambda = 0$. Since the functions $\psi_{\boldsymbol{\nu}}\{\boldsymbol{\nu}\}$ are absolutely arbitrary the sums of integrals of products of the two Q functions multiplying any $\psi_{\boldsymbol{\nu}}$ at any value of the coordinates $\{\boldsymbol{\nu}\}$ must be zero. One thereby obtains integral equation relations between the functions $Q_{n,m}$ for activities 0 and ζ, or, for that matter, between any two other activities.

The equations by themselves would be hopelessly involved, but by making certain omissions tractable expressions can be obtained. The more ambitious of these goes under the name of hyper-netted chains. One involving more drastic-looking approximations is that of Percus–Yevick,† which, however, at least for high temperatures, appears to give rather good numerical results.

6.13. Monte Carlo

No discussion of the statistical mechanical methods of treating dense fluids would be complete without mentioning the type of results now being obtained with the modern high speed computers. The calculations partake of the nature of mathematical experiments, rather than analytical theory, but the results have the enormous advantage that one knows precisely the model of intermolecular forces put into the computation.

Two general methods are used. The one which is the closest to a true experiment is to start, in the computer, a large number, N, of

---

† J. K. Percus and G. J. Yevick, *Phys. Rev.* **110**, 1 (1958).

molecules at given coordinates and velocities and literally compute after successive collisions the positions and velocities. Because of its nature the method is only practical for an assumed hard sphere potential, $u(r) = \infty$, $r \leqslant r_0$, $u(r) = 0, r > r_0$, since only then are single pair collisions instantaneous events uninfluenced by other molecules. This method permits investigation of self-diffusion, and other transport properties as well as deriving equation of state pair distribution function information.

The second method uses a more sophisticated technique. It makes random moves,[†] one molecule at a time. Each move is accepted as a new starting position, or rejected and returned to the original position by a weighting equation so chosen as to obtain, by a statistical average over an infinite number of moves, a distribution proportional to exp $(-U/kT)$. This method is limited to obtaining equilibrium equation of state and distribution function data, but is applicable with realistic mutual potentials. It has, however, up to now, been most used with the hard sphere potentials.

These methods employ a relatively small number, N, of molecules. Periodic boundary conditions are employed, so that a molecule moving out of one wall automatically enters in the same direction at the opposite wall. The validity of the results, despite the relatively small N values, is attested by the fact that the answers are reasonably independent of N between $N = 32$ and N of order 10^3.

Probably the most striking success of these investigations has been to indicate, quite compellingly, that a hard sphere gas will show a first-order transition to a crystal form at something of the order of 1·5 times the close-packed volume. However, in addition, excellent data is available on the pair distribution functions for a hard sphere gas over a very large range of densities.[‡]

[†] The technique of making the computer choose a random number in order to simulate a stochastic process, which, in turn gives the same result as a deterministic equation was first suggested by John von Neumann. He used the term Monte Carlo for this method. Strictly only this second method is Monte Carlo since in the first method only the initial state is random.

[‡] Some references are:
M. N. Rosenbluth and A. W. Rosenbluth, *J. Chem. Phys.* **22**, 881 (1954).
T. Wainwright and B. J. Alder, *Nuovo Cimenta Suppl.* **9**, 116 (1958).
B. J. Alder and T. Wainwright, *J. Chem. Phys.* **27**, 1209 (1957), **31**, 459 (1959).
W. W. Wood and J. D. Jacobson, *J. Chem. Phys.* **27**, 1207 (1957).
W. W. Wood and R. F. Parker, *J. Chem. Phys.* **27**, 720 (1957).

CHAPTER 7

QUANTUM STATISTICS

7.1. Introduction

The original expression of the master equation (1.2.4) for the probability of an open system of fixed V, T, $\boldsymbol{\mu}$, having a given set $\mathbf{N}$ of molecules in a given quantum state $\mathbf{K}$ of energy E_K was given in quantum mechanical language. The examples given, however, have always been to count the quantum states by counting the solutions, $\mathbf{K}^+$, of the Schrödinger equation with numbered molecules, and then to divide this number of states by $\mathbf{N}!$, thus finding the approximate number of truly antisymmetric or symmetric states. This approximation is known under the name "Boltzmann statistics."

In addition to this approximation we have handled the motion of the centers of mass of the molecules by classical equations in all chapters except Chapter 5 on crystals. Now for perfect gases the translational states in any macroscopic volume are indeed so close together in energy that the classical expression is completely valid. It is also true that in the temperature density range for which real gases can be treated as "perfect," the error introduced by using Boltzmann counting is negligible, the corrections necessary to introduce Bose–Einstein or Fermi–Dirac counting never being of larger magnitude than those due to the physical forces between the molecules. However, for imperfect gases of light molecules, H_2, He, and even to some extent Ne, the computation of the virial coefficients should be made using quantum mechanical methods, and, in the case of He at least, a correction for the Bose–Einstein counting is required at low temperatures. For liquids of light molecules at low temperatures quantum methods should certainly be used in any exact treatment. The properties of liquid helium are very different from those to be expected from a hypothetical liquid of classical molecules with the same mutual forces, obeying Boltzmann counting. Even as heavy a molecule as neon shows quantum effects in the liquid, as evidenced by a deviation from the law of corresponding states.

The treatment of the center of mass motion in crystals given in Chapter 5 was quantum mechanical. Although explicit consideration of the requirement that the complete eigenfunctions be symmetric or antisymmetric was not made, the error introduced by this neglect is probably negligible for all real crystals.

The term "quantum statistics" is used for methods in which the corrections for the two effects are made, namely the quantum mechanical treatment of the center of mass motion of real fluids, and the counting of states required for real systems which are always either Bose–Einstein or Fermi–Dirac. As might be expected the two problems are different. The simplest approaches treat them separately. A more sophisticated treatment combines them for those cases where both need to be considered.

The treatment of Bose–Einstein and Fermi–Dirac *perfect* gases is relatively simple, since here we have no problem in describing the correct stationary states, K of the whole gas, and no problem in evaluation of their energies E_K. The treatment given in section 1.5 up to equation (1.5.11) is exact. We need only to avoid the approximation $\mp \ln (1 \mp x) = x$, and treat (1.5.11) more correctly. This is done in sections 7.2 and 7.3 for the two cases respectively. The treatment for the Bose–Einstein gas shows a remarkable condensation phenomenon at low temperature and high densities, which is certainly related to the lambda transition in liquid helium. The treatment is also directly applicable to the photon gas, and leads to the black body radiation equation. The treatment of the Fermi–Dirac gas gives some hint of the behavior to be expected in liquid $\mathrm{He}^{(3)}$, but more important, it gives a useful tool for treating the problem of electrons in metals.

The quantum mechanical treatment of a system of interacting molecules requires a new idea. Our master equation (1.2.4) describes the probability of finding a system in a given stationary state K, of exact energy E_K. These states are to be found by solving the time independent Schrödinger equation $\mathscr{H}\Psi_{K^+} = E_{K^+}\Psi_{K^+}$, with certain boundary conditions. The equation is a differential equation in all of the Γ coordinates, which are some 10^{20} or more in number. Not even a beginning can be made unless the Hamiltonian operator $\mathscr{H}$ is separable into small subsets of coordinates (section 2.1). Now for the perfect gas $\mathscr{H}$ is separable, and a solution is possible. For the crystal we found a way to make $\mathscr{H}$ approximately separable by transforming to normal coordinates expressing small displacements from the equilibrium lattice configuration. The quadratic terms in the development of the potential are then a sum of terms dependent on the single normal coordinates, and as long as the displacements remain small the approximation of neglecting the higher order terms is justifiable. When the crystal melts, however, the displacements are large, and the approximation fails.

The completely general treatment which is discussed in sections 7.5 to 7.10 avoids the necessity of actually finding the stationary states, $\boldsymbol{K}$ and their energies E_K, but uses a moderately sophisticated trick to evaluate the sum $\exp -(E_K/kT)$ for fixed N, without solving the differential equation $\mathscr{H}\Psi_{K^+} = E_{K^+}\Psi_{K^+}$. As used most simply the sum evaluated is that over $\boldsymbol{K}^+$, namely over all solutions of the Schrödinger equation. The simplest use is then to divide by $N!$, since approximately $N!$ different states $\boldsymbol{K}^+$ of the same energy are required to combine into one totally symmetric or totally antisymmetric state of allowed $\boldsymbol{K}$ (section 2.2). However, it is possible to formally write the corrected sum over $\boldsymbol{K}$ and avoid the approximation of Boltzmann counting.

Perhaps one of the most satisfactory results of this technique is to be able to show directly that the classical equations result in the limit $h \rightarrow 0$. These classical equations were arrived at in section 2.3 by using equation (2.3.1) expressing the fact, asserted in that section, that the density of solutions $\boldsymbol{K}^+$ to the Schrödinger equation was always $h^{-\Gamma}$ in the classical phase space. This result follows directly from the technique used in section 7.10.

7.2. Bose–Einstein perfect gases

We choose to start with (1.6.3) which gives the expression,

$$\bar{n}_m = \frac{1}{e^{(\varepsilon_m - \mu)/kT} - 1}, \tag{7.2.1}$$

for the average number, $\bar{n}_m$, of molecules in a given quantum state, $\boldsymbol{m}$, of energy ε_m, for a Bose–Einstein perfect gas. The state $\boldsymbol{m}$ specifies both the internal quantum state and the three translational quantum numbers, k_x, k_y, k_z. The chemical potential, μ, is measured from the zero of energy of the lowest state, which we take to be zero, $\varepsilon_0 = 0$.

From (1) it is immediately evident that if μ approaches zero in value, n_0 approaches infinity. However, $\bar{n}_0$ cannot exceed N, the total number of molecules. If we write $\bar{n}_0 = \alpha N$, $0 < \alpha \leqslant 1$, and solve with $|\mu/kT| << 1$, we find $[\exp(-\mu/kT) - 1] \cong (-\mu/kT) = 1/\alpha N$, namely with N of order 10^{20} as $-\mu/kT$ approaches values of order 10^{-20} an appreciable fraction of all the molecules will be in the lowest quantum state of zero energy. Effectively a value $\mu = -kT/\alpha N$, with α of order unity, is indistinguishable from zero experimentally.

In order to compute the chemical potential μ when it has values less than zero we use for the number, N, of molecules,

$$N = \sum_m \bar{n}_m = \sum_m e^{(\mu-\epsilon_m)/kT}[1 - e^{(\mu-\epsilon_m/kT)}]^{-1}$$

$$= \sum_{n\geqslant 1} \sum_m e^{n(\mu-\epsilon_m)/kT}. \tag{7.2.2}$$

We now restrict ourselves to molecules without internal degrees of freedom, so that the quantum state m is k_x, k_y, k_z of (1.5.2) and the sum over these quantum numbers is given by the integral $(4\pi V/h^3)\int p^2 dp$ with $\epsilon_m = p^2/2m$. One has

$$N/V = \rho = \sum_{n\geqslant 1} e^{n\mu/kT} \int_0^\infty \frac{4\pi}{h^3}\, p^2 e^{-np^2/2mkT}\, dp$$

$$= \sum_{n\geqslant 1} \frac{e^{n\mu/kT}}{\lambda^3 n^{3/2}} \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 e^{-x^2}\, dx$$

where $\lambda^2 = h^2/2\pi mkT$, $x = \sqrt{n}p/\sqrt{2mkT}$.

Since the integral is

$$\int_0^\infty x^2 e^{-x^2}\, dx = \sqrt{\pi}/4,$$

we have

$$\rho = \lambda^{-3} \sum_{n\geqslant 1} n^{-3/2} e^{n\mu/kT}, \tag{7.2.3}$$

in which, for $\rho\lambda^3 << 1$ or $e^{\mu/kT} << 1$ only the first term $n = 1$ need be considered, and this since $\rho = P/kT$, is the expression (1.5.19) for the Boltzmann limit $P/kT = \lambda^{-3}e^{\mu/kT}$.

For larger values of $e^{\mu/kT}$ it is convenient to define dimensionless sums,

$$S_\nu(\mu/kT) = \sum_{n\geqslant 1} n^{\nu-(5/2)} e^{n\mu/kT} \tag{7.2.4}$$

so that

$$dS_\nu/d(\mu/kT) = S_{\nu+1}, \tag{7.2.5}$$

$$(\partial S_\nu/\partial\mu)_T = \frac{1}{kT} S_{\nu+1}, \tag{7.2.6}$$

$$(\partial S_\nu/\partial \ln T)_\mu = -\frac{\mu}{kT} S_{\nu+1}. \tag{7.2.6$'$}$$

These sums all converge for $\mu/kT < 0$, but are singular at $\mu/kT = 0$. The values approached by the sums as μ approaches zero from the negative side are

$$\begin{aligned} \lim_{\mu\to 0}[S_0(\mu/kT)] &= \sum_{n\geq 1} n^{-5/2} = 1.341, \\ \lim_{\mu\to 0}[S_1(\mu/kT)] &= \sum_{n\geq 1} n^{-3/2} = 2.612, \\ \lim_{\mu\to 0}[S_2(\mu/kT)] &= \sum_{n\geq 1} n^{-1/2} = \infty. \end{aligned} \tag{7.2.7}$$

The activity, z, normalized so that $z/\rho = 1$ as both approach zero, is

$$z = \lambda^{-3}e^{\mu/kT}. \tag{7.2.8}$$

If one defines a quantity b_n of dimension $[\text{length}]^{-3(n-1)}$

$$b_1 = 1, \qquad b_n = n^{-5/2}\lambda^{3-3n}, \tag{7.2.9}$$

then

$$\lambda^{-3}S_\nu = \sum_{n\geq 1} n^\nu b_n z^n \tag{7.2.10}$$

completely analogously to the sums of Chapter **4**. We shall later show (section 7.11) that the analogy is not as forced as it appears to be from this definition, but that the b_n are integrals of cluster functions representing a statistical attraction in ordinary space.

We use the thermodynamic relations (1.3.18) and (1.3.19) for a one component system,

$$(\partial P/\partial\mu)_T = N/V = \rho, \tag{7.2.11}$$

$$(\partial P/\partial T)_\mu = S/V. \tag{7.2.12}$$

From (3), which we now write as

$$\rho = \lambda^{-3}S_1(\mu,T), \tag{7.2.13}$$

and (11) we integrate $\rho' d\mu'$ from $\mu' = -\infty$, where $P = 0$, to μ to find from (5) that,

$$\begin{aligned} P &= \int_{-\infty}^{\mu} \rho(\mu' T)d\mu' = kT\lambda^{-3}\int_{-\infty}^{\mu/kT} S_1(x)dx \\ &= kT\lambda^{-3}S_0(\mu/kT). \end{aligned} \tag{7.2.14}$$

With (10) for the sums (13) and (14) are identical in form to (4.2.12) and (4.3.9) for the imperfect gas.

As in Chapter 4, the differentiation of (13) gives an expression for the isothermal compressibility, κ,

$$\kappa = \rho^{-1}(\partial\rho/\partial P)_T = \rho^{-1}(\partial\rho/\partial\mu)_T(\partial P/\partial\mu)_T^{-1},$$

and with (11), (13), and (5),

$$\kappa = \frac{1}{\rho^2 kT}\lambda^{-3} S_2 = \frac{\lambda^3}{kT}\frac{S_2}{S_1{}^2}. \tag{7.2.15}$$

The energy is found by using (12), (14) and (6) with (13), remembering that $d \ln \lambda / d \ln T = -\frac{1}{2}$. One has

$$T(\partial P/\partial T)_\mu = TS/V = \tfrac{5}{2}P - \mu\rho.$$

Since $V\mu\rho = G = PV + E - TS$ one finds the general relationship,

$$E = \tfrac{3}{2}PV, \tag{7.2.16}$$

valid for Boltzmann, Bose–Einstein, and Fermi–Dirac perfect gases, as was proved in section 1.7. For the average energy, ε, per molecule, one then has

$$= \frac{3}{2}\frac{P}{\rho} = \frac{3}{2}kT\frac{S_0}{S_1} \tag{7.2.17}$$

and the heat content per molecule is

$$h = \varepsilon + \frac{P}{\rho} = \frac{5}{2}kT\frac{S_0}{S_1}. \tag{7.2.18}$$

If we use (5) with (13) or (14), remembering that $d \ln \lambda / d \ln T = -\frac{1}{2}$, we can write

$$d\rho = \lambda^{-3}[\tfrac{3}{2}S_1 d \ln T + S_2 d(\mu/kT)],$$

or,

$$dP = kT\lambda^{-3}[\tfrac{5}{2}S_0 d \ln T + S_1 d(\mu/kT)],$$

from which

$$[\partial(\mu/kT)/\partial \ln T]_\rho = -\frac{3}{2}\frac{S_1}{S_2}, \tag{7.2.19}$$

$$[\partial(\mu/kT)/\partial \ln T]_P = -\frac{5}{2}\frac{S_0}{S_1}. \tag{7.2.20}$$

If now these are used in (17) and (18) one has for the heat capacities per molecule,

$$\begin{aligned} C_V = (\partial\varepsilon/\partial T)_\rho &= \frac{3}{2}k\left[\frac{S_0}{S_1} - \frac{3}{2}\frac{S_1}{S_2}\left(1 - \frac{S_2 S_0}{S_1{}^2}\right)\right] \\ &= \frac{15}{4}k\left[\frac{S_0}{S_1} - \frac{3}{5}\frac{S_1}{S_2}\right], \end{aligned} \tag{7.2.21}$$

$$C_P = (\partial h/\partial T)_P = \frac{5}{2}k\left[\frac{S_0}{S_1} - \frac{5}{2}\frac{S_0}{S_1}\left(1 - \frac{S_2 S_0}{S_1^2}\right)\right]$$

$$= \frac{25}{4}k\left[\frac{S_0^2 S_2}{S_1^3} - \frac{3}{5}\frac{S_0}{S_1}\right]. \qquad (7.2.22)$$

Differentiation of (21) with respect to ln T gives

$$(\partial C_V/\partial \ln T)_\rho = -\frac{15}{4}k\frac{3}{2}\frac{S_1}{S_2}\left[1 - \frac{S_0 S_2}{S_1^2} - \frac{3}{5} + \frac{S_1 S_3}{S_2^2}\right]$$

$$= \frac{45}{8}k\left[\frac{S_0}{S_1} - \frac{3}{5}\frac{S_1^2 S_3}{S_2^3} - \frac{2}{5}\frac{S_1}{S_2}\right]. \qquad (7.2.23)$$

Finally the entropy per molecule is $S = (h - \mu/T)$ which, from (18) is,

$$S = k\left[\frac{5}{2}\frac{S_0}{S_1} - \frac{\mu}{kT}\right] = k\left[\frac{5}{2}\frac{S_0}{S_1} - \ln(z\lambda^{-3})\right]. \qquad (7.2.24)$$

Equations (13) to (24) then give expressions for the various thermodynamic functions in terms of the variables μ, T, or with (10) in terms of z, T. To go to the more convenient variables ρ, T one could, in principle, use the β_ν's defined by (4.4.9) in terms of the b_n's, and copy the pertinent equations of Chapter 4. However, the cluster functions (section 7.11) that make up the b_n's of eq. (9) are utterly different from those due to a sum of pair potentials in the classical case, and the β_ν's are not integrals of any simple functions of ν-molecules. Correspondingly they do not have simple numerical coefficients (although β_ν is always proportional to $\lambda^{-3\nu}$). We can, nevertheless, use the numerical method of eq. (4.3.10) with (4.3.11) to find the activity z as a power series in ρ up to any given power, say ρ^4. One can then substitute this in (10) to evaluate the sums. The method is tedious, but straightforward. The results are

$$z = \rho\{1 - 2^{-3/2}(\lambda^3\rho) + [4^{-1} - 3^{-3/2}](\lambda^3\rho)^2$$
$$-[5 \times 2^{-9/2} - 5 \times 2^{-3/2} \times 3^{-3/2} - 2^{-3}](\lambda^3\rho)^3 + \ldots\},$$

$$z = \rho[1 - 0.353\,56(\lambda^3\rho) + 0.057\,55(\lambda^3\rho)^2$$
$$- 0.005\,763(\lambda^3\rho)^3 + \ldots], \qquad (7.2.25)$$

$$P/kT = \rho[1 - 0.176\,78(\lambda^3\rho) - 0.003\,30(\lambda^3\rho)^2$$
$$- 0.000\,110(\lambda^3\rho)^3 - \ldots], \qquad (7.2.26)$$

$$\rho\kappa kT = [1 - 0.353\,56(\lambda^3\rho) - 0.009\,9(\lambda^3\rho)^2$$
$$- 0.000\,44(\lambda^3\rho)^3 - \ldots]^{-1}, \qquad (7.2.27)$$

$$\varepsilon = \tfrac{3}{2}P/\rho, \qquad h = \tfrac{5}{2}P/\rho, \qquad (7.2.28)$$

$$S = k[\tfrac{5}{2} - \ln(\lambda^3\rho) - 0.088\,4(\lambda^3\rho) - 0.003\,33(\lambda^3\rho)^2 - 0.000\,13(\lambda^3\rho)^3 - \ldots], \qquad (7.2.29)$$

and

$$C_V = \tfrac{3}{2}k[1 + 0.088\,39(\lambda^3\rho) + 0.006\,6(\lambda^3\rho)^2 + 0.000\,385(\lambda^3\rho)^3 + \ldots]. \qquad (7.2.30)$$

All of the sums, S_ν, of (4) approach the value $\rho\lambda^3$ for very low densities and high temperatures, $\rho\lambda^3 << 1$. Using this, and the equations (13) to (24), or alternatively with the developments (25) to (30) we see that all thermodynamic functions approach their classical monatomic perfect gas values in this limit, $z = \rho$, $P/kT = \rho$, $\rho\kappa kT = 1$, $\varepsilon = \frac{3}{2}kT$, $S = k[\frac{5}{2} - \ln(\lambda^3\rho)]$ and $C_V = \frac{3}{2}k$. As $\lambda^3\rho$ grows the activity z, the pressure P, the entropy $\bar{S}$, and the average $\bar{\varepsilon}$ are all lower, for fixed ρ, T than the classical values. There are more than the Boltzmann fraction of molecules in the states of zero or near zero kinetic energy, and these, of course, have less energy and exert less pressure than the average. This is the physical reason for the reduced values of P and ε and hence also of z and S. Conversely the heat capacities and compressibility are greater than their classical values.

As temperatures decreases or density increases we approach the singularity in the sums at $\mu = 0$. For larger values of $\rho\lambda^3$, then, μ stays constant at zero ($\mu = -kT/N\alpha$ with $\alpha < 1$ but of order unity) and a non-negligible fraction of the molecules "condense" into the single quantum state of zero energy. The fact that the equations predict this to hold up to infinite density is, of course, a consequence of our treatment of a perfect gas with no forces between molecules, and hence of no repulsion, even at zero distance apart.

This condensation occurs in a line in the plane of T and any one of the variables, μ, z, ρ or P that may be used. The values at the singularity follows from the use of the limiting values of the sums given in (7) used in the equations (13) to (24). The numerical values follow from the numbers in (1.5.22) and (1.5.24). These values are, with M the molecular weight,

$$\left.\begin{aligned}
&\mu_0 = 0 \text{ (independent of } T_0\text{)},\\
&\rho_0 = 2.612\lambda^{-3} = 4.907 \times 10^{20} M^{3/2} T_0^{3/2}\ \text{cm}^{-3}\\
&z_0 = \lambda^{-3} = 1.878 \times 10^{20} M^{3/2} T_0^{3/2}\ \text{cm}^{-3}\\
&z_0\rho_0 = (2.612)^{-1} = 0.383\,2\\
&P_0 = 1.341 kT\lambda^{-3} = 0.034\,32\ M^{3/2} T_0^{5/2}\ \text{atm}\\
&P_0/\rho_0 kT = 1.341/2.612 = 0.513\,4.
\end{aligned}\right\} \qquad (7.2.31)$$

The expression for P_0 is that of the vapor pressure above the condensate in the zero energy state. For He, $M = 4$ the value at the normal boiling point of helium, $T = 4.2$°K is ten atmospheres. The temperatures at which $P_0 = 1$ atm for a perfect gas of molecular weight 4 is $T_0 = 1.32$. Were there no forces between the molecules helium would have these vapor pressures. At the density $\rho = 2.18 \times 10^{22}$ cm^{-3} of liquid helium the value of T_0 computed from (31) is 3.18°K, approximately 50% higher than lambda transition which occurs at 2.19°K.

From (15) that κ is proportional to S_2, and the limit in (7) that $S_2 \rightarrow \infty$ as the singularity is approached one sees that in the condensation of the Bose–Einstein perfect gas, unlike the condensation of a normal classical gas due to molecular forces, the compressibility approaches infinity continuously as the condensation point is approached. This means that there is no break in the slope of the P–V curve, the pressure curve becomes gradually horizontal as V is decreased. Correspondingly the value of C_p from (24) approaches infinity as condensation is reached.

Below the condensation temperature we still retain the relation $\varepsilon = (3/2)P/\rho$. From (32) one sees, then, that for $(\rho\lambda^3) > 2.612$ we have

$$\left.\begin{aligned} \varepsilon &= \tfrac{3}{2}\, 1.341kT(\rho\lambda^3)^{-1} = 2.011kT(\rho\lambda^3)^{-1} \\ C_V &= \tfrac{15}{4}\, 1.341k(\rho\lambda^3)^{-1} = 5.031k(\rho\lambda^3)^{-1} \\ S &= (\varepsilon + P/\rho)T^{-1} = \tfrac{5}{2}\, 1.341k(\rho\lambda^3)^{-1} = 3.3525k(\rho\lambda^3)^{-1}, \end{aligned}\right\} (\rho\lambda^3) > 2.612 \left\{\begin{aligned} &(7.2.32) \\ &(7.2.33) \\ &(7.2.34) \end{aligned}\right.$$

where the last expression $TS = \varepsilon + P/\rho = h$ follows from $\mu = 0$. Since λ is proportional to $T^{-1/2}$ the energy is proportional to $T^{5/2}$ and the other two functions, S and C_V, to $T^{3/2}$. Below the condensation, since the pressure is determined by T alone, no meaning can be assigned to C_p.

The values of these equations may be obtained from (17), (21), and (24) respectively by the following intuitive argument. As the singularity is approached the sums approach the values $S_0 = 1.341$, $S_1 = 2.612$ and $S_2 \rightarrow \infty$, and $\rho\lambda^3 = S_1 = 2.612$. Were the sums extended to $N \cong 10^{23}$ instead of to infinity, as would be appropriate for a finite system, then an extremely small increase in $e^{\mu/KT}$ permits S_1 to keep equal to $\rho\lambda^3$ as this quantity increases. Since the high terms of S_0 are N^{-1} times those in S_1, it follows that S_0 remains constant at $S_0 = 1.341$, whereas S_2 is infinite. If, then, we substitute $S_0 = 1.341$, $S_1 = (\rho\lambda^3)$ and $S_2 = \infty$ in (17), (21), and (24) we do indeed obtain (32), (33) and (34) respectively.

We see from the equations that ε, C_V, S, h as well, of course, as μ and P approach the same values at the condensation from above as from below. There is no discontinuity in these functions. In particular,

$$C_V = \frac{15}{4}\,\frac{1.341}{2.612}\,k = 2.381k \;[\text{at } (\rho\lambda^3) \rightarrow 2.612] \qquad (7.2.35)$$

instead of the classical $C_V = 1.5k$. However, the value of $(\partial C_V/\partial T)_\rho$ can be shown to have a discontinuity at the transition, as we shall proceed to demonstrate.

Since C_V is proportional to $T^{3/2}$ below condensation, we have, from (33),

$$(\partial C_V/\partial \ln T)_\rho = \tfrac{45}{8}\, 1.341k(\rho\lambda^3)^{-1}, \quad [\rho\lambda^3 > 2.612]. \qquad (7.2.36)$$

This is the same value as we would obtain from our intuitive argument based on sums to finite limits from (23). Since S_1 has finite values for the Nth term, and those of S_2 are N-fold larger whereas those of S_3 are N^2-fold larger, we expect S_2 to be of order N and S_3 of order N^2, so that $S_3/S_2{}^3$ will be of order N^{-1} or effectively zero beyond condensation. Only the term S_0S_1 in (23) survives, and with $S_0 = 1.341$, $S_1 = \rho\lambda^3$ this gives (36).

However, as the singularity is approached from above the limit σ of the ratio,

$$\lim_{\alpha\to 0}\left\{\frac{S_3\left(\dfrac{\mu}{kT} = -\alpha\right)}{\left[S_2\left(\dfrac{\mu}{kT} = -\alpha\right)\right]^3}\right\} = \lim_{\alpha\to 0}\left\{\frac{\sum\limits_{n\geq 1} n^{1/2}e^{-\alpha n}}{[\sum\limits_{n\geq 1} n^{-1/2}e^{-\alpha n}]^3}\right\} = \sigma, \quad (7.2.37)$$

will be shown to have the finite value

$$\sigma = 3/2\pi. \qquad (7.2.38)$$

From (23) we then have

$$\begin{aligned}\lim_{\alpha=0}\ (\partial C_V/\partial \ln T)_{\rho,[\rho\lambda^3=2.612-\alpha]} &= \frac{45}{8}\,\frac{1.341}{2.612}\,k - \frac{3}{5}\,\frac{(2.612)^3}{1.341}\,\sigma k\\ &= \frac{45}{8}\,\frac{1.341}{2.612}\,k - \frac{(2.612)^3}{1.341}\,\frac{9}{10\pi}\,k = \frac{45}{8}\,\frac{1.341}{2.612}\,[1 - 3.8]k\\ &= -8.1k, \qquad (7.2.39)\end{aligned}$$

whereas, from (36) the limit approached from the two phase region is $(45/8)(1.341/2.612)k = 2.9k$. The heat capacity, C_V, rises to the cusped maximum of $2.381k$ given by (35) at the transition, and decreases rapidly towards the classical $1.5k$ value above the transition.

The proof of (38) is as follows. Write the sums, with $\Delta n = 1$, as

$$\begin{aligned}\sum_{n>1} n^{\pm 1/2} e^{-\alpha n} &= \sum_{n>1} n^{\pm 1/2} e^{-\alpha n} \Delta n \\ &= \alpha^{\mp(1/2)-1} \sum_{n\geqslant 1} (\alpha n)^{\pm 1/2} e^{-\alpha n} \Delta(\alpha n) \qquad (7.2.40)\end{aligned}$$

where $\Delta(\alpha n) \to 0$ in the limit $\alpha \to 0$. The sums can then be replaced by the integrals, $x = \alpha n$,

$$\int_0^\infty x^{1/2} e^{-x}\, dx = \frac{3}{2}\sqrt{\pi}, \qquad \int_0^\infty x^{-1/2} e^{-x}\, dx = \sqrt{\pi},$$

so that the limit, σ of (37), $\sigma = S_3/S_2{}^3$, is

$$\sigma = \lim_{\alpha \to 0} \frac{\alpha^{-3/2}\, 3\sqrt{\pi}/2}{(\alpha^{-1/2}\sqrt{\pi})^3} = \frac{3}{2\pi}. \qquad (7.2.41)$$

Were we to treat a perfect Bose–Einstein gas with internal degrees of freedom it would, in principle, be necessary to sum (2) independently for each internal state. The only case of some possible practical interest is that of a monatomic gas with spin, having a lowest internal level of zero energy but degeneracy, g, that is consisting of g states. The sum (2) is then the same for each state, and instead of (13) one has $\rho = g\lambda^{-3}S_1$, and the other equations (14) to (24) follow with gS_ν replacing S_ν. The developments (25) to (30) then have $(\lambda^3\rho/g)$ replacing $\lambda^3\rho$ in all terms.

7.3. Black body radiation

The so-called black body radiation consists of the equilibrium photon gas, which is an example of a degenerate Bose–Einstein gas for which $\mu \equiv 0$, differing from that of section (2) only in the differing relation between energy and momentum and in the fact that there are two states of differing polarization for each set k_x, k_y, k_z.

The classical theory of the thermodynamics of black body radiation was of such importance in the development of the quantum theory that at least a brief outline of its salient arguments seems to be in order here.

Consider a space, or "Hohlraum" surrounded by mirror walls which can neither absorb nor emit radiation. Divide this hohlraum into two parts by a partition which reflects, on both sides, all radiation except that between some frequency ν and $\nu + \Delta\nu$, but is transparent in this range of width $\Delta\nu$. At equilibrium the energy density, $u(\nu)\Delta\nu$, of radiation between ν and $\nu + \Delta\nu$ will then be equal on both sides of the partition. Now suppose two small bodies of different materials, maintained at the same temperature T, are placed in the hohlraum on opposite sides of the partition. Each will emit and absorb radiation, and each will tend to produce some characteristic energy density $u(\nu,T)$ per unit frequency range, such that if $u(\nu)$ falls below this value there would be a net absorption of radiation at ν, or if above a net emission. Since, however, the second law requires that at the same temperature there can be no net flow of energy from one body to the other it follows that $u(\nu,T)$ is identical for all bodies, and at every frequency. This function, $u(\nu,T)$, of energy per unit volume per unit frequency range defines the "Black Body Radiation."

As a corollary, if one object is black, absorbing all radiation of frequency ν falling on it, and the other is gray, absorbing only the fraction, $\alpha(\nu) < 1$, then the emissivity of the gray body at frequency ν is $\alpha(\nu)$ times that of the black body. This is one of Kirchoff's Laws.

Now by purely classical electromagnetic theory it is possible to prove that if

$$U(T) = \int_0^\infty u(\nu,T)d\nu \tag{7.3.1}$$

is the total energy per unit volume, then the pressure exerted by the radiation on the walls is given by

$$P(T) = \tfrac{1}{3}U(T). \tag{7.3.2}$$

We have already discussed this at the end of section 1.7 for a thermal equilibrium gas. We will give here a short general proof based on the particle model of photons, each having energy $h\nu$, moving with velocity c, and carrying momentum $h\nu/c$.

A photon striking the mirror wall at angle θ to the normal, and reflected at the same angle, transfers $2 \cos \theta$ times its own momentum to the wall, which, in turn, is $2c^{-1} \cos \theta$ times its energy. In unit time the photons moving in this direction that are in a volume $c \cos \theta$ strike unit of area of wall, so that the momentum per unit time and unit area transferred normal to the wall's surface is $2 \cos^2\theta$ times the energy per unit volume at this angle direction, θ. The momentum transfer per unit area and per unit time is the pressure exerted on the wall by the photons. Since the element of solid angle between θ and

$\theta + d\theta$ is $2\pi \sin \theta \, d\theta$, and the fraction of the total solid angle of 4π is $\frac{1}{2} \sin \theta \, d\theta = \frac{1}{2} d (\cos \theta)$, the ratio of pressure exerted, for each frequency ν, to the energy density at this frequency is

$$P(\nu)/u(\nu) = \int_0^1 \cos^2 \theta d (\cos \theta) = \tfrac{1}{3}$$

for all frequencies. From this and (1) eq. (2) follows.

The next step is a purely thermodynamic proof that $U(T)$ is proportional to T^4, namely that with σ a universal constant,

$$U(T) = \sigma T^4 \tag{7.3.3}$$

The argument considers a Carnot Cycle with a radiation hohlraum as the heat engine, instead of the usual perfect gas. Step (A) of the cycle withdraws the heat Q_1 reversibly and isothermally from a body at T_1, and step (C) returns heat Q_2 reversibly and isothermally to a body at $T_2 = (Q_2/Q_1)T_1$, the intermediate step (B) is the reversible adiabatic expansion of the hohlraum which takes it from T_1 to $T_2 < T_1$. By starting and ending with a zero volume hohlraum the customary fourth step of bringing the gas engine adiabatically back to T_1 is unnecessary, since there is nothing to bring back!

In step (A) we start with a zero volume hohlraum in equilibrium at energy density $U_1 = U(T_1)$ with an (infinite) reservoir at T_1, the pressure being $P_1 = \frac{1}{3} U_1$. Expand the hohlraum reversibly to V_1 doing work $P_1 V_1 = \frac{1}{3} U_1 V_1 = W_1$, and storing the energy $E_1 = U_1 V_1$ in the hohlraum, so that the heat withdrawn is

$$Q_1 = E_1 + W_1 = \tfrac{4}{3} U_1 V_1. \tag{A}$$

Step (B) is the reversible adiabatic expansion of the hohlraum with mirror walls and piston to some volume $V_2 > V_1$ in which process the energy change of the hohlraum is the negative of the work done,

$$dE = d(UV) = UdV + VdU = -PdV = -\tfrac{1}{3} UdV,$$

$$\tfrac{4}{3} d \ln V = -d \ln U,$$

so that

$$V_2/V_1 = (U_1/U_2)^{3/4}. \tag{B}$$

In this step we can ensure that the frequency distribution, $u_2(\nu)$ after completion of the expansion is that of some temperature T_2, $u_2(\nu) = u(\nu, T_2)$, by insertion of a microscopic "black body" of negligible heat capacity which maintains the equilibrium distribution at all times without affecting the energy balance used in deriving (B). Step (C) is now the reverse of step (A), but at this new temperature T_2, for which

$$Q_2 = \tfrac{4}{3}U_2V_2. \tag{C}$$

From (B), with (A) and (C)

$$Q_1/Q_2 = (U_1/U_2)^{1/4}$$

but from the second law, since all steps are reversible,

$$T_1/T_2 = Q_1/Q_2 = (U_1/U_2)^{1/4}$$

and it follows that $U \sim T^4$ so that (3) is derived.

We can now go one step further, and prove that,

$$u(\nu,T) = T^3 f(\nu/T), \tag{7.3.4}$$

where the function f depends only on the one variable, ν/T. Equation (4) is in agreement with (3) since, from (1)

$$U(T) = \sigma T^4 = T^4\int_0^\infty f(\nu/T)d(\nu/T),$$

$$\sigma = \int_0^\infty f(\nu/T)d(\nu/T). \tag{7.3.5}$$

The adiabatic expansion of step (B) preceding cannot be affected by the absence or presence of the microscopic black body which we imagined previously to ensure thermal equilibrium distribution, since, if it were affected we would be able to violate the second law by cycles utilizing our imaginary reflecting partition that is transparent to a particular frequency ν. Now consider the expansion of a cubic mirror walled hohlraum with each edge length L expanding equally, so that $d \ln L = \frac{1}{3}d \ln V$. From B that $V \sim U^{-3/4}$ and since $U \sim T^4$ the expansion is such that $L \sim T^{-1}$. Now each standing wave of the radiation having wavelength λ_1 at edge length L_1 is transformed in the expansion to L_2 to a wave of length $\lambda_2 = \lambda_1(L_2/L_1)$, that is $\lambda \sim T^{-1}$ and the frequency, $\nu = c/\lambda$ is transformed with $\nu \sim T$. All the radiation at T_1 lying between ν_1 and $\nu_1 + \Delta_1\nu$ is transformed to radiation between ν_2 and $\nu_2 + \Delta_2\nu$ with both $\nu_2 = (T_2/T_1)\nu_1$ and $\Delta_2\nu = (T_2/T_1)\Delta_1\nu$.

Since the total energy density varies as T^4

$$u(\nu_2 T_2)\Delta_2\nu = (T_2/T_1)^4 u(\nu_1 T_1)\Delta_1\nu$$

we have eq. (4), which is known as the Wien displacement law.

Returning now to the overall thermodynamic quantity, $U(T)$, the energy per unit volume, $U(T) = \sigma T^4$ we see that the heat capacity, C_V, per unit volume is

$$C_V = dU/dT = 4\sigma T^3, \tag{7.3.6}$$

and the entropy density, S, is

$$S = \int_0^T C_V d \ln T = \tfrac{4}{3}\sigma T^3. \tag{7.3.7}$$

The Gibbs free energy, G, per unit volume is $G = U + P - TS$, or, from (2), (3) and (7),

$$G = 0. \tag{7.3.8}$$

These rather far reaching conclusions, summarized by eqs. (3), (2), (6), (7) and (8) for the thermodynamic functions, as well as (4) for the frequency distribution of the energy density, had all been arrived at in the late nineteenth century. Then came the disaster!

The standing radiation waves of the mirror walled cubic Hohlraum of edge L are those of component wavelengths $\lambda_x = 2L/k_x, \lambda_y = 2L/k_y$, $\lambda_z = 2L/k_z$ with k_x, k_y, k_z positive integers. The frequency, ν, is

$$\nu = c/\lambda,$$

where

$$\frac{1}{\lambda} = \sqrt{\frac{1}{\lambda_x^2} + \frac{1}{\lambda_y^2} + \frac{1}{\lambda_z^2}} = \frac{k}{2L}$$

$$k = \sqrt{k_x^2 + k_y^2 + k_z^2}.$$

For sufficiently large values of k, between k and $k + dk$ there are $4\pi k^2 dk/8$ different positive integral values, k_x, k_y, k_z, and since each standing wave of given k_x, k_y, k_z can have either of two planes of polarization there are

$$C(V,k)dk = \pi k^2 dk = (8\pi V/c^3)\nu^2 d\nu \tag{7.3.9}$$

different standing waves of the hohlraum in the frequency interval between ν and $\nu + d\nu$. Now each standing wave is an oscillator, and the classical equilibrium average energy of an oscillator at temperature T is kT so that

$$Vu(\nu,T)d\nu = (8\pi VkT/c^3)\nu^2 d\nu,$$

$$u(\nu,T) = (8\pi k/c^3)T^3(\nu/T)^2 \tag{7.3.10}$$

in agreement with (4). But the disaster is that σ, from (5) has the value infinity. The energy density per unit frequency range should classically increase without limit as ν^2 increases.

Empirically eq. (10), the Rayleigh–Jeans distribution, was found to hold for sufficiently small values of ν/T. An empirical law, the Wien distribution law,

$$u(\nu,T) = (8\pi/c^3)h\nu^3 e^{-h\nu/kT} \tag{7.3.10'}$$

(which also agrees with (4)) and which contained a parameter h of dimension energy × time, was found to be valid at high $h\nu/kT$ values. Planck interpolated the equation

$$u(\nu,T) = (8\pi/c^3)\frac{h\nu^3}{e^{h\nu/kT}-1} \tag{7.3.11}$$

which becomes (10) for $h\nu/kT << 1$, and (10′) for $h\nu/kT >> 1$, and which fits the intermediate range. It was then Planck's genius to recognize that the existence of this law required a fundamental alteration in our classical laws, and to state the nature of this alteration, namely that radiation could only be emitted or absorbed in units of $h\nu$.

We may derive the Planck law simply by replacing the classical average energy, kT, of an oscillator by its quantum mechanical value

$$kT \rightarrow \frac{h\nu}{e^{h\nu/kT}-1}$$

and multiplying by $(8\pi/c^3)\nu^2$ for the number of independent oscillators per unit volume and frequency range.

A derivation treating the photons as a perfect gas of photons of energy $h\nu$, and chemical potential $\mu = 0$ was given by the Indian physicist Bose in 1924, and sent to Einstein, who in turn communicated it with a foreword emphasizing its importance to the Prussian Academy. In this paper Bose first introduced what has been called the Bose–Einstein Statistics, namely the correct logical counting of state of a system composed of "Bosons," or particles for which the many body eigenfunction is symmetric in permutations.

In our development we obtain the equivalent of the Bose derivation by noting that since, from (8) $G = 0$ we have $\mu = 0$ in the photon gas. We then use (7.2.1)

$$\bar{n}_m = \frac{1}{e^{h\nu/kT}-1}$$

for the number of photons in each quantum state $m = k_x, k_y, k_z$ and fixed polarization. Multiply this by $h\nu$, the energy of the photon, and $(8\pi/c^3)\nu^2$ for the number of states per unit volume and frequency range, and one has (11). Thus the black body radiation is merely an example of an equilibrium gas of degenerate Bose–Einstein particles with $\mu = 0$, and with the number of states per unit energy range given by $8\pi(hc)^{-3}\varepsilon^2$.

If (11) is written as

$$u(\nu,T)d\nu = T^4 \frac{8\pi k^4}{(hc)^3} \frac{(h\nu/kT)^3}{e^{h\nu/kT} - 1} d\left(\frac{h\nu}{kT}\right)$$

one has, with (1) and (3),

$$\int_0^\infty u(\nu,T)d\nu = U(T) = \sigma T^4$$

$$\sigma = \frac{8\pi k^4}{(hc)^3} \int_0^\infty \frac{x^3dx}{e^x - 1}. \tag{7.3.12}$$

The definite integral is

$$\frac{x^3dx}{e^x - 1} = \sum_{n\geqslant 1} \int_0^\infty x^3 e^{-nx}dx = \left(\sum_{n\geqslant 1} \frac{1}{n^4}\right) \int_0^\infty y^3 e^{-y}dy$$

$$= 6 \sum_{n\geqslant 1} \frac{1}{n^4} = 6\,\frac{\pi^4}{90}.$$

From this and (12)

$$\sigma = \frac{24}{45}\,\pi^5 k^4 (hc)^{-3}.$$

7.4. Fermi–Dirac perfect gases

Equation (1.6.3) for a Fermi–Dirac system gives, for the average number, $\bar{n}_m$, of molecules in a state m of energy ε_m,

$$\bar{n}_m = \frac{1}{e^{(\varepsilon_m - \mu)/kT} + 1} \tag{7.4.1}$$

instead of (7.2.1) for the Bose–Einstein gas, in which the denominator is the exponential minus unity. Since the exponential is always positive the value of $\bar{n}_m$ in the Fermi–Dirac system is always less than unity, as is required, since for these systems no two molecules are allowed to have the same quantum state.

Qualitatively, the behavior of the gas differs from that of a Boltzmann gas in the opposite direction from the Bose–Einstein deviations. At low temperature and high density, where the deviations are largest there are fewer molecules in the low energy states than in the Boltzmann case. As a result the energy and pressure are higher than classical, the heat capacity lower, since as the temperature rises the difference, F–D minus Boltzmann, decreases. There is nothing resembling the Bose–Einstein condensation. The molecules behave as if they repelled one another.

As long as μ is negative, which we shall see later is not generally the case, we could develop the sum as in (7.2.2) and obtain

$$\begin{aligned} N &= \sum_m \bar{n}_m = \sum_m e^{(\mu-\varepsilon_m)/kT}[1 + e^{(\mu-\varepsilon_m)/kT}]^{-1} \\ &= \sum_{n\geqslant 1} (-)^{n-1} \sum_m e^{n(\mu-\varepsilon_m)/kT} \end{aligned} \qquad (7.4.2)$$

which differs from (7.2.2) only in the alternation of the sign with increasing n. The same series that occur in section 7.2 then appear, but with alternating signs. The series diverge at $\mu = 0$, $z = \lambda^{-3}$, but the divergence is due to singularities in the functions they represent at imaginary μ-values, $\mu/kT = i\pi(2n + 1)$ for which $e^{\mu/kT} = -1$, $z_0 = -\lambda^{-3}$. These are unallowed values for the physical system, and the functions are actually analytic for the physically allowed μ-real, z-positive real range, through the divergence of the series. The series can be used, and are useful for high temperatures and low densities, but the more interesting region where μ is positive needs another mathematical technique.

The most interesting physical case which approximates, at least, a perfect Fermi–Dirac gas is that of electrons in metals. The valence electrons, one, two, or three per atom, depending on the column in the periodic system, can be regarded as "free." This arises from the fact that the electrons move throughout the metal in the electric field of the residual positive ions, and since the coulombic potential is the long range r^{-1} type it is not very strongly dependent on position, and can, in zero'th order approximation, be regarded as constant within the metal, but with an abrupt rise at the surface. Thus this approximation leads to a perfect gas model, from which one may start various perturbation calculations to take into account the periodic nature of the lattice and the interactions of the electrons with each other and with the lattice vibrations of the ions. We develop here only the perfect gas case.

Electrons have half unit spins, and there are $g = 2$ states, $m_s = \pm \frac{1}{2}$, for each translational state. The number of states between energy ε and $\varepsilon + d\varepsilon$ is then g times $C_{\mathrm{tr}}(V,\varepsilon)d\varepsilon$ of (1.7.6) or

$$C(V,\varepsilon)d\varepsilon = 2\pi g V(2m)^{3/2}h^{-3}\varepsilon^{1/2}d\varepsilon. \qquad (7.4.2')$$

Since the mass, m, of the electron is so small this density of states is some 10^{-5} fold less, per unit energy range, than it would be for a gas of molecular weight unity. The total number $M(V,\varepsilon_0)$ of states below an energy ε_0 is

$$M(V,\varepsilon_0) = \int_0^{\varepsilon_0} C(V,\varepsilon)d\varepsilon = (4\pi g/3)V(2m\varepsilon_0/h^2)^{3/2} \qquad (7.4.3)$$

$$M(V,\varepsilon_0)/V = (4g/3\sqrt{\pi})(2\pi m\varepsilon_0/h^2)^{3/2}. \qquad (7.4.3')$$

At absolute zero the electrons will occupy the lowest allowed states, one per state or $g = 2$ per translational state, up to some energy ε_0, determined by the number density ρ, as

$$(2\pi m\varepsilon_0/h^2) = (3\sqrt{\pi}\rho/8)^{2/3}$$

$$\varepsilon_0/k = (h^2/2\pi mk)(3\sqrt{\pi}/8)^{2/3}\rho^{2/3}.$$

In eqs. (1.5.21) and (22) we found $h^2/2\pi mk \simeq 300/M(\text{ÅU})^2$ deg. or with $M \simeq 0.5 \times 10^{-3}$ for electrons $(h^2/2\pi m_e k) \simeq 6 \times 10^5(\text{ÅU})^2$ deg. Even for metals such as cesium with high molecular volumes and only one valence electron, ρ is of order $10^{-2}(\text{ÅU})^{-3}$ and ε_0/k is many thousand degrees. At zero temperature, then, the Pauli exclusion principle requires that the electrons occupy states up to kinetic energies ε_0, whose energy is that of the classical kT at temperatures of order 10^4 deg. This energy,

$$\varepsilon_0 = (6/\pi g)^{2/3}(h^2\rho^{2/3}/8m), \qquad (7.4.4)$$

and the corresponding momentum,

$$p_0 = \sqrt{2m\varepsilon_0} = \tfrac{1}{2}(6/\pi g)^{1/3}h\rho^{1/3}, \qquad (7.4.5)$$

are known as the energy and momentum, respectively, of the "Fermi surface," the surface of the spherical "Fermi sea" in momentum space, within which all states are occupied by particles and above which there are none, at $T = 0°\text{K}$.

One may now anticipate the results shown by the subsequent equations. As T increases there appear "holes" or unfilled states to a depth of order kT in the Fermi sea, and states in the atmosphere above the sea become occupied to a similar height. Since relatively few states are in this narrow energy band the heat capacity is low, and the thermodynamic functions are relatively little changed with temperature.

Turning to the number, $\bar{n}_m$ of electrons (or molecules) per state, eq. (1), we see that in the limit $T \rightarrow 0$, the exponential in the denominator approaches zero for $\varepsilon_m - \mu < 0$ and infinity for $\varepsilon_m - \mu > 0$, so that $\bar{n}_m$ is unity for $\varepsilon_m < \mu$ and zero for $\varepsilon_m > \mu$. It follows that at $T = 0$,

$$\mu(T = 0) = \varepsilon_0. \tag{7.4.6}$$

For any T, using (1) and (2), integrating over ε to find N and dividing by V,

$$\rho = N/V = (4\pi g\sqrt{2})(m/h^2)^{3/2}\int_0^\infty \frac{\varepsilon^{1/2}d\varepsilon}{e^{(\varepsilon-\mu)/kT} + 1}. \tag{7.4.7}$$

Multiply both sides by $(\varepsilon_0{}^{3/2}/\rho)$ with (4) for ε_0 to find

$$\varepsilon_0{}^{3/2} = \frac{3}{2}\int_0^\infty \frac{\varepsilon^{1/2}d\varepsilon}{e^{(\varepsilon-\mu)/kT} + 1} \tag{7.4.8}$$

as an implicit integral equation for μ at any temperature.

The integration of (8) and subsequent solution of the resulting implicit numerical equation for μ as ε_0 times unity plus a power series in the small quantity $(KT/\varepsilon_0)^2$ is comparatively long and tricky, but not difficult. It is advantageous to simplify the notation by introducing

$$\begin{aligned} y &= \mu/kT, & y_0 &= \varepsilon_0/kT \\ \zeta &= (\varepsilon - \mu)/kT, & \varepsilon &= kT(y + \zeta) \end{aligned} \tag{7.4.9}$$

so that, dividing (8) by $(kT)^{3/2}$ on both sides, one has

$$\int_{-y}^\infty \frac{1}{e^\zeta + 1}\left[\tfrac{3}{2}(y + \zeta)^{1/2}\right]d\zeta = y_0{}^{3/2} \tag{7.4.10}$$

to solve for y as a function of y_0.

Although for the purposes of this section we need only solve (10), there arise more general problems in the use of Fermi statistics in which one wishes to solve for y a general form of equation,

$$\int_{-y}^\infty n(\zeta)g(y + \zeta)d\zeta = G(y_0), \tag{7.4.10$'$}$$

with $y_0 >> 1$, in which

$$n(\zeta) = (1 + e^\zeta)^{-1}, \tag{7.4.11}$$

$$n'(\zeta) = dn/d\zeta = -e^\zeta(1 + e^\zeta)^{-2} = -(e^{\zeta/2} + e^{-\zeta/2})^{-2}, \tag{7.4.11$'$}$$

and

$$G(y_0) = \int_0^{y_0} g(x)dx. \tag{7.4.12}$$

We make use of the fact that $n(\zeta)$ is nearly a step function, unity for $\zeta << 0$ and zero for $\zeta >> 1$, so that its derivative, $n'(\zeta)$ of (11′) is even in ζ with value $-1/4$ at $\zeta = 0$, decreasing exponentially to zero on both sides of zero. Since $G = 0$ at one limit of the integration in (10′), namely at $\zeta = -y$, and $n = 0$ at the other, $\zeta = \infty$, the product $Gn = 0$ at both limits, and the integrand in (10′) can be replaced by $-Gn'$. Now $n'(\zeta)$ is of order e^{-y} at the lower limit, with $y \simeq y_0$, and therefore negligible. With error only of the order e^{-y_0} we can therefore extend the lower limit to $\zeta = -\infty$, and try to solve for y in

$$\int_{-\infty}^{+\infty} G(y + \zeta)n'(\zeta)d\zeta = G(y_0). \tag{7.4.13}$$

We now assume G to be a slowly varying function of ζ at $\zeta \simeq 0$, so that the quantities

$$\lambda_\nu = y^\nu[G(y)]^{-1}d^\nu G/dy^\nu = y^\nu G^{-1}d^{\nu-1}g/dy^{\nu-1} \tag{7.4.14}$$

are of the same order of magnitude for successive ν-values. For our case of eq. (10),

$$G(y) = y^{\frac{3}{2}}, \qquad g(y) = \tfrac{3}{2}y^{\frac{1}{2}}, \tag{7.4.15}$$

$$\lambda_1 = \tfrac{3}{2}, \qquad \lambda_2 = \tfrac{3}{4}, \qquad \lambda_3 = -\tfrac{3}{8}, \qquad \lambda_4 = \tfrac{9}{16},$$

and the λ_ν's are pure numbers, independent of y, which they would be if G were any power of y.

We then use the development

$$G(y + \zeta) = G(y)\left[1 + \sum_{\nu \geqslant 1} \frac{\lambda_\nu}{\nu\,!y^\nu}\zeta^\nu\right] \tag{7.4.16}$$

in (13). The integral of the first term, that of $-n'(\zeta)d\zeta$, is unity from (11) and (11′). Since n' is even in ζ the integrals of the odd powers of ζ vanish, and only the integrals,

$$I_n = -\frac{1}{(2n)!}\int_{-\infty}^{\infty} \zeta^{2n}n'(\zeta)d\zeta = -\frac{2}{(2n\,!)}\int_0^{\infty} \zeta^{2n}n'(\zeta)d\zeta \tag{7.4.17}$$

survive. With (16) and (17) in (13) we have

$$G(y)\,[1 + \sum_{n \geqslant 1} (\lambda_{2n}I_n)y^{-2n}] = G(y_0) \tag{7.4.18}$$

which is now an implicit numerical equation for y in terms of y_0.

For sufficiently large y_0 values the solution of (18) is $y = y_0$. We try in (18) a development up to y_0^{-4} of the form

$$y = y_0[1 - \alpha y_0^{-2} - \beta y_0^{-4} - \ldots], \tag{7.4.19}$$

since evidently only even negative powers appear. With this, using $y^{-2} = y_0^{-2} + 2\alpha y_0^{-4} + \ldots$, we have,

$$\begin{aligned} 1 + \lambda_2 I_1 y^{-2} + \lambda_4 I_2 y^{-4} + \ldots = 1 &+ \lambda_2 I_1 y_0^{-2} \\ &+ (\lambda_4 I_2 + 2\alpha\lambda_2 I_1) y_0^{-4} + \ldots \end{aligned}$$

and, developing $G(y)$ about $G(y_0)$ with the analogue of (16),

$$\begin{aligned} G(y) &= G(y_0)\left[1 + \lambda_1 \frac{y - y_0}{y_0} + \frac{1}{2}\lambda_2 \left(\frac{y - y_0}{y_0}\right)^2 + \ldots\right] \\ &= G(y_0)[1 - \alpha\lambda_1 y_0^{-2} + (\tfrac{1}{2}\alpha^2\lambda_2 - \beta\lambda_1) y_0^{-4} - \ldots] \end{aligned}$$

so that (18) becomes

$$\begin{aligned} 1 + (\lambda_2 I_1 - \alpha\lambda_1) y_0^{-2} + (\lambda_4 I_2 &+ 2\alpha\lambda_2 I_1 - \alpha\lambda_1\lambda_2 I_1 \\ &+ \tfrac{1}{2}\alpha^2\lambda_2 - \beta\lambda_1) y_0^{-4} + \ldots = 1, \end{aligned}$$

which is satisfied by

$$\begin{aligned} \alpha &= (\lambda_2/\lambda_1) I_1, \\ \beta &= (\lambda_4/\lambda_1) I_2 + (\lambda_2/\lambda_1)^2 I_1^2 [2 - \lambda_1 + \tfrac{1}{2}(\lambda_2/\lambda_1)]. \end{aligned} \tag{7.4.20}$$

The next step is to find the values of the I_n's of (17), with n' given by (11′). We write

$$-n' = e^{-\zeta}(1 + e^{-\zeta})^2 = \sum_{\nu \geqslant 1} (-)^{\nu-1} \nu e^{-\nu\zeta}$$

so that

$$\begin{aligned} I_n &= 2 \sum_{\nu \geqslant 1} (-)^{\nu-1} \nu^{-2n} \left[\int_0^\infty \frac{(\nu\zeta)^{2n}}{(2n)!} e^{-\nu\zeta} d(\nu\zeta)\right] \\ &= 2 \sum_{\nu \geqslant 1} (-)^{\nu-1} \nu^{-2n}. \end{aligned} \tag{7.4.21}$$

The Bernoulli numbers are

$$\begin{aligned} B_n &= \sum_{\nu \geqslant 1} \nu^{-2n} \\ B_1 = \pi^2/6, &\qquad B_2 = \pi^4/90. \end{aligned} \tag{7.4.22}$$

The sums in (21) alternate in sign. They can be written as the sum over all ν with positive sign minus twice the sum over the even values. But the sums over even ν's are 2^{-2n} times the sums over all ν-values. Hence

$$I_n = 2[1 - 2^{-(2n-1)}]B_n$$
$$I_1 = \pi^2/6 \qquad I_2 = 7\pi^4/360. \tag{7.4.23}$$

With these values in (20) for α and β, in eq. (19),

$$y = y_0\, 1 - \frac{1}{6}\left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{\pi}{y_0}\right)^2 - \frac{1}{360}\left[7\left(\frac{\lambda_4}{\lambda_1}\right) + 10\left(\frac{\lambda_2}{\lambda_1}\right)^2\left(2 - \lambda_1 + \frac{\lambda_2}{2\lambda_1}\right)\right]\frac{\pi^4}{y_0} + \ldots. \tag{7.4.24}$$

Equation (24) then gives a solution of (10′) for (y/y_0) up to terms y_0^{-4} where the y_ν's are the quantities defined by (14). For our particular case (10) the values of λ_ν are those listed in (15), and with y and y_0 given by (9) we have

$$y = y_0\left[1 - \frac{\pi^2}{12}\, y_0^{-2} - \frac{\pi^4}{80}\, y_0^{-4} - \ldots\right], \tag{7.4.25}$$

as the solution of (10) for $y = \mu/kT$ in terms of $y_0 = \varepsilon_0/kT$, an expansion which is valid for sufficiently high values of y_0.

We now seek expressions for the other dimensionless energy quantities divided by kT, namely for ε/kT, $P/\rho kT$, and $Ts/kT = s/k$, where ε is the average energy per particle and s the entropy per particle. We may well expect as is actually the case, that these will turn out to be functions of y_0 alone. We have already proven, section 1.7, that for any perfect gas,

$$\varepsilon = \tfrac{3}{2}P/\rho \tag{7.4.26}$$

so that if we define,

$$x = P/\rho kT, \tag{7.4.27}$$

we have

$$\varepsilon/kT = \tfrac{3}{2}x \tag{7.4.28}$$

and

$$y = \mu/kT = [\varepsilon + (P/\rho) - Ts]/kT = \tfrac{5}{2}x - (s/k). \tag{7.4.29}$$

The intensive properties of a system are fully determined by ρ and T, but since $y_0 = \varepsilon_0/kT$ and ε_0 varies as $\rho^{2/3}$ from (4), we can, instead use y_0 and T, with

$$(\partial \ln \rho/\partial \ln T)y_0 = \tfrac{3}{2} \quad (\partial \ln \rho/\partial y_0)_T = 3/2y_0, \tag{7.4.30}$$

so that, from (27), $P = \rho kTx$, we have

$$\frac{1}{\rho kT}\,dP = \left[\left(\frac{\partial x}{\partial y_0}\right)_T + \frac{3}{2}\frac{x}{y_0}\right]dy_0 + \left[\left(\frac{\partial x}{\partial \ln T)}\right)_{y_0} + \frac{5}{2}x\right]d\ln T. \tag{7.4.31}$$

Use this in the familiar

$$d\mu = \rho^{-1}dP - Ts\,d\ln T,$$

$$dy = d\left(\frac{\mu}{kT}\right) = \frac{1}{\rho kT}\,dP - \frac{\mu + Ts}{kT}\,d\ln T,$$

with (29) to find,

$$dy = \left[\left(\frac{\partial x}{\partial y_0}\right)_T + \frac{3}{2}\frac{x}{y_0}\right]dy_0 + \left(\frac{\partial x}{\partial \ln T}\right)_{y_0} d\ln T. \tag{7.4.32}$$

Since y depends only on y_0 we have

$$(\partial y/\partial \ln T)y_0 = (\partial x/\partial \ln T)y_0 = 0,$$

and hence x is a function of y_0 alone, as expected.

We have from (32)

$$\frac{dx}{dy_0} + \frac{3}{2}\frac{x}{y_0} = {y_0}^{-3/2}\frac{d}{dy_0}\,{y_0}^{3/2}\,x(y_0) = \frac{dy(y_0)}{dy_0},$$

or

$$x(y_0) = {y_0}^{-3/2}\int^{V_0} z^{3/2}\,\frac{dy(z)}{dz}\,dz$$

with some, as yet unknown, constant of integration. Using (25), which is valid for high z values,

$$z^{3/2}dy/dz = z^{3/2} + (\pi^2/12)z^{-1/2} + (3\pi^4/80)z^{-5/2},$$

the integration gives

$$x = x_0 + \frac{2}{5}\,y_0\left[1 + \frac{5\pi^2}{12}\,{y_0}^{-2} - \frac{\pi^4}{16}\,{y_0}^{-4} - \ldots\right]. \tag{7.4.33}$$

With eqs. (27) to (29), and (25) we may now list the thermodynamic equations,

$$\mu = \varepsilon_0\left[1 - \frac{1}{12}\,(\pi kT/\varepsilon_0)^2 - \frac{1}{80}\,(\pi kT/\varepsilon_0)^4 - \ldots\right], \tag{7.4.34}$$

$$\varepsilon = \frac{3}{5}\varepsilon_0\left[1 + \frac{5}{12}(\pi kT/\varepsilon_0)^2 - \frac{1}{16}(\pi kT/\varepsilon_0)^4 + \ldots\right], \tag{7.4.35}$$

$$P = (2/3)\varepsilon, \tag{7.4.36}$$

$$a = \mu - (P/\rho) = \frac{3}{5}\varepsilon_0\left[1 - \frac{5}{12}(\pi kT/\varepsilon_0)^2 + \frac{1}{48}(\pi kT/\varepsilon_0)^4 + \ldots\right] \tag{7.4.37}$$

$$s = -(\partial a/\partial T) = \frac{1}{2}(\varepsilon_0/T)\left[(\pi kT/\varepsilon_0)^2 - \frac{1}{20}(\pi kT/\varepsilon_0)^4 + \ldots\right]. \tag{7.4.38}$$

7.5. Density matrix formulation

The mechanical state of a single *classical* system is described by giving a point, $\boldsymbol{p}^{(\Gamma)}\boldsymbol{q}^{(\Gamma)}$

$$\boldsymbol{p} = \boldsymbol{p}^{(\Gamma)} = p_1, p_2, \ldots, p_\alpha, \ldots p^\Gamma \tag{7.5.1}$$

$$\boldsymbol{q} = \boldsymbol{q}^{(\Gamma)} = q_1, q_2, \ldots, q_\alpha, \ldots q^\Gamma \tag{7.5.1'}$$

in the Γ-space of 2Γ variables, the Γ coordinates, q_α, $1 \leqslant \alpha \leqslant \Gamma$, and their Γ conjugated momenta, p_α. The dimensions of any coordinate times that of the conjugated momentum is always $q_\alpha p_\alpha \sim ml^2t^{-1}$, which are the dimensions of Planck's constant h.

An infinite ensemble of classical systems can then be fully defined by giving a probability density function, $W_{\text{Cl}}(\boldsymbol{pq})$ in this Γ-space, which function gives the probability density of finding a member system of the ensemble at the point $\boldsymbol{p},\boldsymbol{q}$. The function W_{Cl} is necessarily real and positive everywhere, and normalized so that

$$\iint W(\boldsymbol{p},\boldsymbol{q})d\boldsymbol{p}d\boldsymbol{q} = 1. \tag{7.5.2}$$

The average values, $\bar{f}$, of any function $f(\boldsymbol{p}^{(\Gamma)}\boldsymbol{q}^{(\Gamma)})$ of the coordinates and momenta is then calculable by the equation

$$\bar{f} = \int f(\boldsymbol{p},\boldsymbol{q})W(\boldsymbol{p},\boldsymbol{q})d\boldsymbol{p}d\boldsymbol{q}. \tag{7.5.3}$$

The mechanical state of a single quantum mechanical system may be defined by giving a normalized state function, $\Phi_i(\boldsymbol{q})$ (which may be complex) in the coordinate space $\boldsymbol{q}$, or alternatively, for the same state i, the related normalized state function, $\chi_i(\boldsymbol{p})$ in the momentum space $\boldsymbol{p}$. The normalization of these functions requires that, if $\Phi_i^*(\boldsymbol{q})$ is the conjugate complex of Φ_i, the integral

$$\int \Phi_i(\boldsymbol{q})\Phi_i^*(\boldsymbol{q})d\boldsymbol{q} \equiv \Phi_i \cdot \Phi_i^* = 1, \tag{7.5.4}$$

where the dot product notation is a convenient shorthand representation of the integral of the product of two functions.

The function Φ_i must obey certain conditions. It must be single valued, differentiable, and square integrable so that (4) can be satisfied. It must obey boundary conditions of being zero outside the limits of the system. It must be symmetrical in the exchange of pairs of identical Bose molecules and antisymmetrical in exchange of identical Fermi molecules. It need *not* be a solution of the time independent Schrödinger equation,

$$\mathscr{H}\Psi_K = E_K\Psi_K, \tag{7.5.5}$$

with $\mathscr{H}$ the Hamiltonian operator. *If* the state is a solution of (5) it is a stationary state, namely such that the properties of the system do not change with time. More generally the state, Φ_i, obeys the time dependent Schrödinger equation,

$$\partial\Phi_i/\partial t = (2\pi i/h)\mathscr{H}\Phi_i, \tag{7.5.6}$$

which, in principle, determines the state for future time, just as the classical equations of motion determine the future of a classical system for which $\boldsymbol{p},\boldsymbol{q}$ is given at time t.

An ensemble of M quantum mechanical systems, having each the same set $\boldsymbol{N} = N_1, N_2, \ldots$ of molecules in the same volume V, can then be described if the states, $\Phi_i(\boldsymbol{q})$ or $\chi_i(\boldsymbol{p})$, are known for each system i, $1 < i < M$. The mathematical function which is most convenient to represent such an ensemble is the density matrix, $\boldsymbol{\Gamma}$, for the ensemble, and it is this density matrix which plays the same role in determining the average value, $\bar{f}$, of any function $f(\boldsymbol{p},\boldsymbol{q})$ of momenta and coordinates that is played by the probability density function for an ensemble of classical systems. The density matrix, $\boldsymbol{\Gamma}$, may be given in a variety of representations. Different representations of a matrix correspond vaguely to various choices of variable transformations for an ordinary function. In the coordinate representation the elements $\Gamma(\boldsymbol{q}',\boldsymbol{q}'')$ of the matrix have continuous row variables $\boldsymbol{q}'$ and continuous column variables $\boldsymbol{q}''$, so that the element is a function of 2Γ single variables, just as is the classical probability density function, $W_{\text{Cl}}(\boldsymbol{p},\boldsymbol{q})$. A particular representation, the Wigner representation, $W(\boldsymbol{p},\boldsymbol{q})$, in coordinate and momentum space corresponds most closely to the classical function $W_{\text{Cl}}(\boldsymbol{p},\boldsymbol{q})$.

We return, first, to the discussion of a single system i in state $\Phi_i(q)$. The physical interpretation of the state is that its absolute value square, $\Phi_i(\boldsymbol{q})\Phi_i^*(\boldsymbol{q}) = |\Phi_i(\boldsymbol{q})|^2$ is the probability density that the single system i will be found at the coordinate q, and hence the average value $\bar{f}_i$ of any function, $f(q)$, for this system is given by

$$\bar{f}_i = \int \Phi_i^*(q) f(q) \Phi_i(q) dq. \tag{7.5.7}$$

This average is called the "expectation value" of $f(q)$ in the state i. The algorithm for finding the expectation value $\bar{f}_i$ of a function $f(p,q)$ of coordinates and momenta is as follows. In $f(p,q)$ replace each momentum p_α by $(h/2\pi i)\partial/\partial q_\alpha$. This forms an operator

$$\mathscr{F}(p,q) = f[(h/2\pi i)\nabla_q, q]. \tag{7.5.8}$$

The expectation value $\bar{f}_i$ is then

$$\bar{f}_i = \int \Phi_i^*(q) \mathscr{F}(p,q) \Phi_i(q) dq. \tag{7.5.9}$$

In particular if f is a product of powers of the momenta (and we note that any analytic function $f(p)$ can be expanded as a sum of powers),

$$f(p) = p^n = \prod_\alpha p_\alpha^{n_\alpha}, \tag{7.5.10}$$

then the operator $\mathscr{F}(p)$ is,

$$\mathscr{F}(p) = \mathscr{P}^n = \prod_\alpha (h/2\pi i)^{n_\alpha} (\partial^{n_\alpha}/\partial q_\alpha^{n_\alpha}), \tag{7.5.11}$$

and

$$\langle p^n \rangle_{\text{Av}} = \int \Phi_i^*(q) \mathscr{P}^n \Phi_i(q) dq. \tag{7.5.12}$$

There is one remark to be made about the general operator. If, in $f(p,q)$ there occurs a product $p_\alpha q_\alpha$, it makes no difference in the classical case whether one writes this as $q_\alpha p_\alpha$ or as $p_\alpha q_\alpha$. The quantum mechanical average is different for the two cases, since the operators are

$$\begin{aligned} q_\alpha p_\alpha &\rightarrow (h/2\pi i) q_\alpha (\partial/\partial q_\alpha), \\ p_\alpha q_\alpha &\rightarrow (h/2\pi i)(\partial/\partial q_\alpha) q_\alpha \\ &= h/2\pi i + (h/2\pi i) q_\alpha (\partial/\partial q_\alpha). \end{aligned}$$

From the normalization (4) the use of these in (9) gives the well-known commutation relation,

$$\langle p_\alpha q_\alpha - q_\alpha p_\alpha) \rangle_{\text{Av}} = h/2\pi i.$$

The expression (9) for finding the expectation value, $\bar{f}_i$, can be written in a shorter notation, and one which is more suitable for use with ensembles. For two sets, q' and q'' of values of the coordinates define a rather trivial product function,

$$\Gamma_i(q', q'') = \Phi_i^*(q') \Phi \ \ (q'') \tag{7.5.13}$$

which is a function of 2Γ variables. This function is conjugate symmetric in the exchange of $\boldsymbol{q}'$ and $\boldsymbol{q}''$,

$$\Gamma_i(\boldsymbol{q}',\boldsymbol{q}'') = \Gamma_i^*(\boldsymbol{q}'',\boldsymbol{q}'). \tag{7.5.14}$$

Regard $\Gamma_i(\boldsymbol{q}',\boldsymbol{q}'')$ as the $\boldsymbol{q}',\boldsymbol{q}''$ element, in the coordinate representation of a Hermitian matrix, Γ_i, where $\boldsymbol{q}'$ is the continuous row variable and $\boldsymbol{q}''$ the column variable. The diagonal elements, $\Gamma_i(\boldsymbol{q},\boldsymbol{q})$ with $\boldsymbol{q}' = \boldsymbol{q}'' = \boldsymbol{q}$, are real positive, and are the probability densities in the coordinate space. The "trace" of the matrix is the sum of the diagonal elements, or in this case where the element variable is continuous it is the integral, and is unity† from (4),

$$[\Gamma_i]_{\text{tr}} = \int \Gamma_i(\boldsymbol{q},\boldsymbol{q})d\boldsymbol{q} = 1. \tag{7.5.15}$$

This matrix, Γ_i, is the "density matrix" for the system in the state Φ_i.

Let $\boldsymbol{F}_i$ be a matrix of elements $F_i(\boldsymbol{q}',\boldsymbol{q}'')$ which are formed by letting the operator $\mathscr{F}(\boldsymbol{p}',\boldsymbol{q}')$ operate on the variable $\boldsymbol{q}'$ *only* of $\Gamma_i(\boldsymbol{q}',\boldsymbol{q}'')$,

$$F_i(\boldsymbol{q}',\boldsymbol{q}'') = \mathscr{F}(\boldsymbol{p}',\boldsymbol{q}')\Gamma_i(\boldsymbol{q}',\boldsymbol{q}''). \tag{7.5.16}$$

Equation (9) for finding the expectation value $\bar{f}_i$ in state i is now

$$\bar{f} = [F_i]_{\text{tr}}. \tag{7.5.17}$$

The product AB of two matrices A and B, which is different from BA, is a matrix, the x,y element of which is $A(x,z)B(z,y)$ summed or integrated over the middle index z depending on whether z is discrete or continuous respectively. The product $\Gamma_i\Gamma_i$, then, has, from (13) and (4) elements that are

$$\begin{aligned}\Gamma_i^2(\boldsymbol{q}',\boldsymbol{q}'') &= \int \Gamma_i(\boldsymbol{q}',\boldsymbol{q}''')\Gamma_i(\boldsymbol{q}''', \boldsymbol{q}'')d\boldsymbol{q}''' \\ &= \Gamma_i(q',q'').\end{aligned}$$

In matrix notation we have

$$\boldsymbol{\Gamma}_i\boldsymbol{\Gamma}_i = \boldsymbol{\Gamma}_i, \tag{7.5.18}$$

i.e. the density matrix for a single system is equal to its own square. This property is called "itempotent."

We later have use for another notation for the operations symbolized by (16) and (17). We utilize the Dirac delta function, which, for a single variable y is defined to be a function, $\delta(y)$, which is zero everywhere

---

† Various normalizations other than unity are often used. We use (15) to emphasize the analogy to a probability density.

except at $y = 0$ and there of such (infinite) magnitude that its integral over y is unity,

$$\int_{-\infty}^{+\infty} \delta(y)dy = 1. \tag{7.5.19}$$

This very singular function has the convenient property that if $f(y)$ in any well behaved function of y that

$$\int_{-\infty}^{+\infty} f(y)\delta(y - x)dy = f(x). \tag{7.5.20}$$

Introducing

$$D(\boldsymbol{q}' - \boldsymbol{q}'') = \prod_{\alpha=1}^{\alpha=\Gamma} \delta(q_\alpha' - q_\alpha'') \tag{7.5.21}$$

as the product over all coordinates α of their Dirac delta functions, one may then write the trace condition, (15), as

$$[\Gamma_i]_{tr} = \iint D(\boldsymbol{q}' - \boldsymbol{q}'')\Gamma_i(\boldsymbol{q}',\boldsymbol{q}'')d\boldsymbol{q}'d\boldsymbol{q}'' = 1, \tag{7.5.22}$$

and the expression for obtaining the expectation value, $\bar{f}_i$, as

$$\bar{f}_i = \iint D(q' - q'')\mathscr{F}(p',q')\Gamma_i(q',q'')dq'dq''. \tag{7.5.23}$$

The density matrix $\boldsymbol{\Gamma}$ of an ensemble may now be defined. The ensemble consists of M systems, labeled by the running index i, $1 \leqslant i \leqslant M$, and the state, $\Phi_i(\boldsymbol{q})$ and therefore the density matrix, $\boldsymbol{\Gamma}_i$, of the system i for every i-value is assumed to be known. The density matrix, $\boldsymbol{\Gamma}$, of the ensemble is then defined as the average of all $\boldsymbol{\Gamma}_i$,

$$\boldsymbol{\Gamma} = M^{-1} \sum_{i=1}^{i=M} \boldsymbol{\Gamma}_i, \tag{7.5.24}$$

so that, from (13), its elements are

$$\Gamma(\boldsymbol{q}',\boldsymbol{q}'') = M^{-1} \sum_{i=1}^{i=M} \Phi_i(\boldsymbol{q}')\Phi_i^*(\boldsymbol{q}''). \tag{7.5.24'}$$

The normalization is that the trace is unity, from (22) and (24),

$$[\Gamma]_{\text{tr}} = \iint \Gamma(\boldsymbol{q}',\boldsymbol{q}'')D(\boldsymbol{q}' - \boldsymbol{q}'')d\boldsymbol{q}'d\boldsymbol{q}'' = 1. \tag{7.5.25}$$

The average value, $\bar{f}$, of a function $f(\boldsymbol{p},\boldsymbol{q})$ is the average of the expectation values $\bar{f}_i$,

$$\bar{f} = M^{-1}\Sigma\bar{f}_i = [F]_{\text{tr}}, \tag{7.5.26}$$

$$F(\boldsymbol{q}',\boldsymbol{q}'') = \mathscr{F}(\boldsymbol{p}',\boldsymbol{q}')\Gamma(\boldsymbol{q}',\boldsymbol{q}''). \tag{7.5.26'}$$

Two explanatory comments are appropriate here. Firstly, had we attempted to define an averaged normalized wave function,

$$\Phi(q) = A\sum_{i}\Phi_i(q),$$

with A so chosen that $\Phi \cdot \Phi^* = 1$, and from this $\Phi(\boldsymbol{q})$ construct a matrix of elements $\Phi(\boldsymbol{q}')\Phi^*(\boldsymbol{q}'')$ the matrix so defined would be different from (24) in that it would contain additive non-zero contributions, $A^2\Phi_i(q')\Phi_j^*(q'')$, with $i \neq j$ from *different systems*. These terms, in turn, would give improper contributions to $\bar{f}$, were this matrix used in (26), and $\bar{f}$ would no longer be the average of the expectation values of the independent systems.

The second comment concerns complete "orthonormal" sets of functions. Functions, $\Theta_{\boldsymbol{M}}(\boldsymbol{q})$ are said to be normalized if, for each $\boldsymbol{M}$-value $\Theta_{\boldsymbol{M}} \cdot \Theta_{\boldsymbol{M}}^* = 1$, and to be mutually orthogonal if $\Theta_{\boldsymbol{M}'} \cdot \Theta_{\boldsymbol{M}''}^* = 0$ for $\boldsymbol{M}' = \boldsymbol{M}''$. The combined property is written as,

$$\Theta_{\boldsymbol{M}'} \cdot \Theta_{\boldsymbol{M}''}^* = \delta(\boldsymbol{M}' - \boldsymbol{M}''), \tag{7.5.27}$$

with $\delta(\boldsymbol{M}' - \boldsymbol{M}'')$, called the Kronecker delta symbol, having values,

$$\delta(\boldsymbol{M}' - \boldsymbol{M}'') = 1 \text{ if } \boldsymbol{M}' = \boldsymbol{M}'',$$

$$\delta(\boldsymbol{M}' - \boldsymbol{M}'') = 0 \text{ if } \boldsymbol{M}' \neq \boldsymbol{M}''. \tag{7.5.27'}$$

A set of functions obeying (27) is said to be a set of orthonormal functions. The set is complete if it contains an infinite number of members, $0 \leqslant \boldsymbol{M} \leqslant \infty$, which may all obey certain boundary and symmetry conditions, such that *any* function, $\Phi_i(\boldsymbol{q})$, obeying these conditions can be expressed as an (infinite) sum of the set of $\Theta_{\boldsymbol{M}}$'s,

$$\Phi_i(\boldsymbol{q}) = \sum_{\boldsymbol{M}} a_{i\boldsymbol{M}}\Theta_{\boldsymbol{M}}(\boldsymbol{q}), \tag{7.5.28}$$

where, from (27), the coefficients $a_{i\boldsymbol{M}}$ are,

$$a_{i\boldsymbol{M}} = \Phi_i \cdot \Theta_{\boldsymbol{M}}^*. \tag{7.5.28'}$$

The totality of all normalized solutions, Ψ_K, of the time independent Schrödinger equation (5) form such a complete orthogonal set, and so do the solutions of (5) if some fictitious approximate Hamiltonian $\mathscr{H}^{(0)}$ of the same number of coordinates is used. It is often convenient to use such a complete set to express arbitrary functions by the relation (28), and the state functions Φ_i of the ith system can then be defined by giving the numbers $a_{i\boldsymbol{M}}$ for some arbitrary complete orthonormal set $\Theta_{\boldsymbol{M}}$.

The state function Φ_i is always normalized by (4), so that, from (27), using (28),

$$\sum_M a_{iM} a_{iM}^* = 1 \tag{7.5.29}$$

and the positive product $a_{iM} a_{iM}^*$ is called the probability or expectation value that the system i will, if analysed, be found in the arbitrarily defined state Θ_M. It is *not* assumed, however, that two *different systems* i and j need be in orthogonal states, that is $\Phi_i \cdot \Phi_j^*$ with $i \neq j$ need *not* be zero, so that $\sum\limits_M a_{iM} a_{jM}^*$ need not be zero.

If (28) is used the density matrix Γ_i of the *system* i has elements,

$$\Gamma_i(q',q'') = \sum_{M'} \sum_{M''} a_{iM'} a_{iM''}^* \Theta_{M'}(q') \Theta_{M''}^*(q''), \tag{7.5.30}$$

including "off diagonal" terms with $M' \neq M''$, and that of the ensemble has also,

$$\Gamma(q',q'') = \sum_{M'} \sum_{M''} \Theta_{M'}(q') \Gamma_{M',M''} \Theta_{M''}(q'') \tag{7.5.31}$$

with

$$\Gamma_{M',M''} = M^{-1} \sum_{i=1}^{i=M} a_{iM'} a_{iM''}^*, \tag{7.5.32}$$

which is not necessarily zero for $M' \not\equiv M''$. The numbers $\Gamma_{M',M''}$ are called the M',M'' elements of the matrix $\mathbf{\Gamma}$ in the representation of the quantum number, M describing the discrete orthonormal complete set Θ_M. The diagonal elements, $\Gamma_{M,M}$, which are a sum of absolute values squared from (32), are all positive. From (29) and (32) their sum is unity. This is a particular example of the general rule that the trace formed by summing or integrating the diagonal elements of a matrix is invariant under the representation,

$$[\Gamma]_{\text{tr}} = \sum_M \Gamma_{M,M} = 1. \tag{7.5.33}$$

Whereas the diagonal elements, $\Gamma_{M,M}$ are necessarily positive, the product $a_{iM'} a_{iM''}^*$ for $M' = M''$ will generally be complex, and may as well have negative as positive real or imaginary parts. It is often assumed that the element of the ensemble $\Gamma_{M',M''}$ of (32) for $M' \not\equiv M''$, is negligible, and that in the limit of an ensemble of an infinite number of members, $M \to \infty$, these off diagonal terms become zero. This assumption is actually made in constructing the density matrix for an equilibrium system (section 7.9).

7.6. The momentum representation

The state of a single system may be defined by a normalized state function, $\Phi_i(\boldsymbol{q})$, expressed as a function of the coordinates $\boldsymbol{q}$, or alternatively by a normalized function $\chi_i(\boldsymbol{p})$ expressed as a function of the momenta. Actually one could also prescribe a function of q_α for some degrees of freedom, and the momenta p_β for others. For the same state, the functions $\Phi_i(\boldsymbol{q})$ and $\chi_i(\boldsymbol{p})$ are Fourier transforms of each other.

For a single variable, y, limited to the range $-\frac{1}{2}l \leqslant y \leqslant \frac{1}{2}l$, the functions of integer n values, $-\infty < n < \infty$,

$$\phi_n(y) = (1/\sqrt{l})e^{2\pi i n y/l}, \quad \phi_n{}^* = \phi_{-n}, \tag{7.6.1}$$

form a complete orthonormal set,

$$\phi_n \cdot \phi_m{}^* = \delta(n - m) \tag{7.6.1'}$$

and any function, $f(y)$, defined in this range can be developed as the infinite Fourier sum,

$$f(y) = \sum_{n=-\infty}^{n=+\infty} a_n \phi_n(y), \qquad a_n = f \cdot \phi_n{}^*. \tag{7.6.2}$$

In the limit that the range, l, goes to infinity, the sum of (2) goes over into the Fourier integral. One uses the variable,

$$k = n/l, \qquad n = kl, \tag{7.6.3}$$

where the allowed values of k become continuous in the limit. The sum over n goes over to an integral

$$\sum_n \to \int dn = l \int dk,$$

$$f(y) = \int l a_{(kl)} \phi_{(kl)}(y) dk = \int [l^{1/2} a_{(kl)}] e^{2\pi i k y} dk,$$

or with $g(k) = l^{1/2} a_{(kl)}$ we have

$$f(y) = \int_{-\infty}^{+\infty} g(k) e^{2\pi i k y} dk, \tag{7.6.4}$$

$$g(k) = \int_{-\infty}^{+\infty} f(y) e^{-2\pi i k y} dy. \tag{7.6.4'}$$

The orthonormality condition (1') used for the functions $e^{2\pi i k y} = l^{1/2} a_{(kl)}$ now goes over into a Dirac delta function of the variable $k - k'$. This can be seen as follows. Define $\delta(k - k')$ as the integral

$$\int_{-\infty}^{+\infty} e^{2\pi i (k - k') y} dy = \delta(k - k'). \tag{7.6.5}$$

From (1′) it is seen that when the limits are the finite limits of $-\frac{1}{2}l$ and $\frac{1}{2}l$ the function $\delta(k - k')$ is zero when $k - k' \neq 0$, and has the value l when $k - k' = 0$. The allowed values of k are $k = n/l$ so that in the range dk there are dk/l allowed values, and the integral $\int \delta(k - k')dk'$ has the value unity in the limit, and when the range of integration includes $k - k' = 0$. Thus the function defined by (5) has the two defining properties of the Dirac delta function: it is non-zero only at $k - k' = 0$, and its integral over k' is unity.

The nature of the coordinates, q_α, for the various α values has not been specified, and any convenient set could be used. Among others, angular coordinates limited to ranges such as $-\pi \leqslant q_\alpha \leqslant \pi$ might occur. For such coordinates the discrete Fourier transform, (2), rather than the integral (4) should be used, and the "allowed" values of the subsequently introduced momenta, p_α, would be the allowed values for angular momenta, $p_\alpha = n_\alpha h/2\pi$. If, however, the coordinates are cartesian their range can be taken as infinite, even though the value of $\Phi_i(q)$ is necessarily zero for values outside of the boundaries of the system. In this case the Fourier integral form is called for. We will use the notation suitable to this. No fundamental difficulty would be introduced by using discrete summation for some, and integration for other coordinates, and the conclusions would be unaltered, but the notation is awkward.

Instead of writing the transforms in terms of the variables k_α above we introduce, for each coordinate α, a variable p_α related to k_α by $p_\alpha = hk_\alpha$. The dimensions of Planck's constant h are ml^2t^{-1} which are always those of a coordinate multiplied by its conjugate momentum. Since k_α has the dimension reciprocal to the coordinate q_α the quantity p_α has then always the dimension of the momentum conjugate to the coordinate α. As before let p symbolize a set of p_α's for every α and define the symbol $\boldsymbol{p} \cdot \boldsymbol{q}$ to mean

$$\boldsymbol{p} \cdot \boldsymbol{q} = \Sigma p_\alpha q_\alpha. \tag{7.6.6}$$

The function

$$e^{(2\pi i/h)\boldsymbol{p} \cdot \boldsymbol{q}} = \prod_{\alpha=1}^{\alpha=\Gamma} e^{(2\pi i/h)p_\alpha q_\alpha} \tag{7.6.7}$$

is now used for the complete transform, where, from (5), since $dp_\alpha = hdk_\alpha$,

$$h^{-\Gamma} \int_{-\infty}^{+\infty} e^{(2\pi i/h)(q'' - q') \cdot p} dp = D(\boldsymbol{q}' - \boldsymbol{q}), \tag{7.6.8}$$

and Γ, the number of degrees of freedom, is the total number of coordinates. The symmetry of $\boldsymbol{p}$ and $\boldsymbol{q}$ also gives,

$$h^{-\Gamma}\int_{-\infty}^{+\infty} e^{(2\pi i/h)(\boldsymbol{p}'-\boldsymbol{p}'')\cdot\boldsymbol{q}}\,d\boldsymbol{q} = D(\boldsymbol{p}'-\boldsymbol{q}'') \tag{7.6.8$'$}$$

The functions $\chi_i(\boldsymbol{p})$ and $\Phi_i(\boldsymbol{q})$ are related by the Fourier transforms,

$$\chi_i(\boldsymbol{p}) = h^{-1/2\Gamma}\int \Phi_i(\boldsymbol{q})e^{(-2\pi i/h)\boldsymbol{p}\cdot\boldsymbol{q}}\,d\boldsymbol{q}, \tag{7.6.9}$$

$$\Phi_i(\boldsymbol{q}) = h^{-1/2\Gamma}\int \chi_i(\boldsymbol{p})e^{(2\pi i/h)\boldsymbol{p}\cdot\boldsymbol{q}}\,d\boldsymbol{p}. \tag{7.6.9$'$}$$

We shall show that the function $\chi_i(\boldsymbol{p})$ defined by (9) has actually, for the momentum variables $\boldsymbol{p}$, the corresponding interpretation and usefulness that $\Phi_i(\boldsymbol{q})$ has for the coordinate variables.

In the first place we note that χ_i is normalized, namely from (9) with (8),

$$\begin{aligned}\chi\cdot\chi_i^* &= h^{-\Gamma}\iiint\Phi_i(\boldsymbol{q}')\Phi_i^*(\boldsymbol{q}'')e^{(2\pi i/h)\boldsymbol{p}\cdot(\boldsymbol{q}''-\boldsymbol{q}')}\,d\boldsymbol{p}d\boldsymbol{q}'d\boldsymbol{q}''\\ &= \iint\Phi_i(\boldsymbol{q}')\Phi_i^*(\boldsymbol{q}'')D(\boldsymbol{q}''-\boldsymbol{q}')d\boldsymbol{q}'d\boldsymbol{q}'' = 1.\end{aligned} \tag{7.6.10}$$

The interpretation that the absolute value square, $\chi_i(\boldsymbol{p})\chi_i^*(\boldsymbol{p})$, should be the probability density of finding the system i at the momentum value $\boldsymbol{p}$ requires that for any function, $f(\boldsymbol{p})$, its expectation value, $\bar{f}_i$, for the state i should be given by the classical integral,

$$\bar{f} = \int\chi_i(\boldsymbol{p})f(\boldsymbol{p})\chi_i^*(\boldsymbol{p})d\boldsymbol{p}, \tag{7.6.11}$$

and in particular for a power such as that of (7.5.10)

$$\langle \boldsymbol{p}^n\rangle_{\mathrm{Av}} = \int\chi_i(\boldsymbol{p})\mathscr{P}^n\chi_i^*(\boldsymbol{p})d\boldsymbol{p}. \tag{7.6.11$'$}$$

Conversely if (11′) is correct for every power n, then (11) holds for every analytic function f, and if (11) holds, then $\chi_i(\boldsymbol{p})\chi_i^*(\boldsymbol{p})$ *is* the probability density. But using the method (7.5.11) and (7.5.12) prescribed for computing $\langle \boldsymbol{p}^n\rangle_{\mathrm{Av}}$ from Φ_i with (9′) we have

$$\begin{aligned}\langle \boldsymbol{p}^n\rangle_{\mathrm{A}} &= h^{-\Gamma}\iiint\chi_i^*(\boldsymbol{p}')e^{-(2\pi i/h)\boldsymbol{p}'\cdot\boldsymbol{q}}\chi_i(\boldsymbol{p})\mathscr{P}^n e^{(2\pi i/h)\boldsymbol{p}\cdot\boldsymbol{q}}\,d\boldsymbol{p}'d\boldsymbol{p}d\boldsymbol{q}\\ &= h^{-\Gamma}\iiint\chi_i^*(\boldsymbol{p}')\boldsymbol{p}^n\chi_i(\boldsymbol{p})e^{(2\pi i/h)\boldsymbol{q}\cdot(\boldsymbol{p}-\boldsymbol{p}')}\,d\boldsymbol{p}'d\boldsymbol{p}d\boldsymbol{q}\\ &= \iint\chi_i^*(\boldsymbol{p}')\boldsymbol{p}^n\chi_i(\boldsymbol{p})\,D(\boldsymbol{p}-\boldsymbol{p}')\,d\boldsymbol{p}d\boldsymbol{p}',\\ &= \int\chi_i^*(\boldsymbol{p})\boldsymbol{p}^n\chi_i(\boldsymbol{p})d\boldsymbol{p},\end{aligned}$$

which is the same as (11′).

One may similarly show that for computing a function of $\boldsymbol{q}$ from χ_i the rule is to replace q_α by the operator $-(h/2\pi i)\partial/\partial p_\alpha$.

The function $\Gamma(\boldsymbol{p}',\boldsymbol{p}'')$ analogous to $\Gamma(\boldsymbol{q}',\boldsymbol{q}')$

$$\Gamma(\boldsymbol{p}',\boldsymbol{p}'') = M^{-1}\sum_{i=1}^{i=M}\chi_i(\boldsymbol{p}')\chi_i{}^*(\boldsymbol{p}'') \tag{7.6.12}$$

is the $\boldsymbol{p}',\boldsymbol{p}''$ element of the density matrix, $\boldsymbol{\Gamma}$ in the momentum representation. From (7.4.24′) and (9) or (9′) the two representations are related by

$$\Gamma(\boldsymbol{p}',\boldsymbol{p}'') = h^{-\Gamma}\iint e^{-(2\pi i/h)\boldsymbol{p}'\cdot\boldsymbol{q}'}\,\Gamma(\boldsymbol{q}'\boldsymbol{q}')e^{(2\pi i/h)\boldsymbol{q}''\boldsymbol{p}''}\,d\boldsymbol{q}'d\boldsymbol{q}'' \tag{7.6.13}$$

$$\Gamma(\boldsymbol{q}',\boldsymbol{q}'') = h^{-\Gamma}\iint e^{(2\pi i/h)\boldsymbol{q}'\cdot\boldsymbol{p}'}\,\Gamma(\boldsymbol{p}',\boldsymbol{p}'')e^{-(2\pi i/h)\boldsymbol{p}''\cdot\boldsymbol{q}''}\,d\boldsymbol{p}'d\boldsymbol{p}''. \tag{7.6.13′}$$

7.7. Matrix transformations

The matrix transform, (7.6.13), and its inverse, (7.6.13′), is a special case of a unitary transformation of a Hermitian matrix, and of the inverse transform. Hermitian matrices, such as the density matrix $\boldsymbol{\Gamma}$, or matrices representing physical entities such as the energy, may be represented by elements such as $\Gamma(\boldsymbol{q}',\boldsymbol{q}'')$, $\Gamma(\boldsymbol{p}',\boldsymbol{p}'')$ or $\Gamma_{M',M''}$ of (7.5.32) in which the row and column variables may be either continuous, such as $\boldsymbol{q}',\boldsymbol{q}''$, or discrete indices, as in the case of M',M''. Representations in which both the row and column index have the same physical significance, i.e. both are coordinates, or both are momenta, or, both are quantum number designations, M', M'' of the same complete set of orthonormal functions, are most generally used.

The transformation from one representation, say y, to another, x, is done by the use of a unitary matrix, U. The matrix U has element $u(x,y)$ where, in physical examples, x and y are different in meaning: for instance one may be the value of a quantum number, and may only take discrete values, whereas the other may be the continuous momentum or coordinate variable, or perhaps the discrete set of quantum numbers representing a different complete orthonormal set of functions. With the matrix U there is associated another matrix, $U\dagger$, called its conjugate transpose, which has elements

$$u\dagger(y,x) = u^*(x,y). \tag{7.7.1}$$

The matrix U is said to be unitary if, and only if,

$$\begin{aligned}
&\left.\begin{array}{ll}\int u(y',x)u\dagger(x,y'')dx & (x \text{ continuous})\\ \sum\limits_x u(y',x)u\dagger(x,y'') & (x \text{ discrete})\end{array}\right\} = \delta(y'-y''),\\
&\left.\begin{array}{ll}\int u\dagger(x',y)u(y,x'')dy & (y \text{ continuous})\\ \sum\limits_y u\dagger(x,y)u(y,x) & (y \text{ discrete})\end{array}\right\} = \delta(x'-x''),
\end{aligned} \tag{7.7.2}$$

and the δ symbol stands for the Kronecker delta if the variable is discrete, or for the Dirac function if the variable is continuous. The simpler matrix notation is,

$$\begin{aligned} UU^* &= \mathbf{1}, \\ U^*U &= \mathbf{1}, \end{aligned} \qquad (7.7.2')$$

where $\mathbf{1}$ is the symbol for the "unit matrix," namely a matrix for which all elements are zero except the diagonal elements, and these are δ-functions: i.e. the Kronecker delta unity, for each diagonal element when the rows and columns have discrete indices, or the Dirac delta function when the variables are continuous. The rule of summing or integrating over the middle index to obtain the elements of the product matrix is then such that for any matrix, A, the product with the unit matrix is A, i.e.

$$\mathbf{1}A = A\mathbf{1} = A. \qquad (7.7.3)$$

The transformation of a matrix A in the y representation, A_y, to the x representation, A_x is accomplished by

$$A_x = UA_yU\dagger, \qquad (7.7.4)$$

meaning that the element $A(x',x'')$ is

$$A(x',x'') = \int\int u(x',y')A(y',y'')u\dagger(y'',x'')dy'dy''. \qquad (7.7.4')$$

From (2′), (3) and (4) the inverse transform is

$$U\dagger A_x U = U\dagger UA_yU\dagger U = \mathbf{1}A_y\mathbf{1} = A_y, \qquad (7.7.5)$$

$$A(y',y'') = \int\int u\dagger(y',x')A(x',x'')u(x'',y'')dx'dx''. \qquad (7.7.5')$$

If either y or x is a discrete index, rather than a continuous variable, the integration operation of (4′) or (5′) is to be read as a double summation.

The functions, $\theta_n(y)$, of any complete orthonormal set may be regarded as the y,n elements of a unitary matrix $\mathbf{\Theta}$, the conjugate transpose of which, $\mathbf{\Theta}\dagger$, has n,y elements which are $\theta_n{}^*(y)$ from (1). The relation

$$\mathbf{\Theta}\dagger\mathbf{\Theta} = \mathbf{1} \qquad (7.7.6)$$

is then the usual orthonormality condition,

$$\int\theta_n{}^*(y)\theta_m(y)dy = \delta(n-m), \qquad (7.7.6')$$

whereas the relation

$$\mathbf{\Theta}\mathbf{\Theta}\dagger = \mathbf{1} \qquad (7.7.7)$$

is a less well-known relation that, for the infinite sum of all functions,

$$\sum_n \theta_n(y')\theta_n{}^*(y'') = \delta(y' - y''). \tag{7.7.7'}$$

For the density matrix, $\mathbf{\Gamma}_q$, in the representation $\Gamma(q',q'')$ of the Γ-dimensional coordinate q, the transform to a (Γ-dimensional) quantum number representation $\mathbf{\Gamma}_m$ of a particular complete orthonormal set of functions $\Theta_M(\mathbf{q})$ is written as

$$\mathbf{\Gamma}_m = \mathbf{\Theta}\dagger\mathbf{\Gamma}_q\mathbf{\Theta}, \tag{7.7.8}$$

which means that the elements are

$$\Gamma_{M'M''} = \iint\Theta_{M'}{}^*(\mathbf{q}')\Gamma(\mathbf{q}',\mathbf{q}'')\Theta_{M''}(\mathbf{q}'')d\mathbf{q}'d\mathbf{q}'', \tag{7.7.8'}$$

which, from (7.5.24′) and (7.5.28′), is the same as (7.5.32). The inverse transformation

$$\mathbf{\Gamma}_q = \mathbf{\Theta}\mathbf{\Gamma}_m\mathbf{\Theta}\dagger \tag{7.7.9}$$

has the elements given by eq. (7.5.31).

The matrix $U_{p,q}$ of elements

$$u(\mathbf{p},\mathbf{q}) = h^{-1/2\Gamma}\, e^{-(2\pi i/h)\mathbf{p}\cdot\mathbf{q}} \tag{7.7.10}$$

has a conjugate transpose $U_{qp} = U_{pq}\dagger$ of elements

$$u\dagger(\mathbf{q},\mathbf{p}) = h^{-1/2\Gamma}\, e^{(2\pi i/h)\mathbf{q}\cdot\mathbf{p}}, \tag{7.7.10'}$$

and is seen from (7.6.8) and (7.6.8′) to be unitary. This is the matrix that transforms from the Γ-dimensional $\mathbf{q}$-space of infinite range to the continuous Γ-dimensional $\mathbf{p}$-space. As mentioned before, if some of the coordinates, q_α, have finite range, as do angle variables, these p_α-values are discrete. The transformations (7.6.13) and (7.6.13′) are then transforms of the type

$$\mathbf{\Gamma}_p = U_{pq}\mathbf{\Gamma}_q U_{qp}, \tag{7.7.11}$$

$$\mathbf{\Gamma}_q = U_{qp}\mathbf{\Gamma}_p U_{pq}. \tag{7.7.11'}$$

One comment about notation may be made here. It is customary to refer to the matrix of a physical entity, such as the energy given by the Hamiltonian operator, as the Hamiltonian matrix, independently of its representation. Similarly, the density matrix $\mathbf{\Gamma}$ of a particular ensemble is regarded as the same matrix whatever the representation. In this language a unitary transformation does not change the matrix. We have, in this section, written a subscript on the matrix, $A_x, A_y, \mathbf{\Gamma}_p$, $\mathbf{\Gamma}_q, \mathbf{\Gamma}_m$, etc., to indicate the representation when writing the matrix equations for the unitary transformation, (4), (5), (8), (9), (11), (11′), since without some such designation the equation would be meaningless.

A very useful theorem exists which states that any Hermitian matrix can always be brought to diagonal form. This means that, for the density matrix, $\mathbf{\Gamma}$, as example, there is always some set $\Theta_M(q)$ of orthonormal functions in the coordinate space $\boldsymbol{q}$, such that $\Gamma_{M',M''}$ of (8′) has only diagonal elements, $M' \equiv M''$, all others being zero. For the density matrix these diagonal elements are necessarily all positive.

Another comment may be made here. The relation (7) or (7′) for any complete orthonormal set of functions means that for an ensemble in which every one of the infinite set of functions $\Theta_M(q)$ that diagonalize $\mathbf{\Gamma}$ were to occur with equal probability, $\Gamma_{M,M} = a$, the coordinate representation, and indeed any representation, would have diagonal elements only. Now such a density matrix corresponds to no physically realizable ensemble, since it implies equal probabilities of all states of all energies up to infinity. However, it is to be expected that if very many states are represented with not greatly differing probabilities the coordinate and momenta representations will be strongly peaked at the diagonal, and will approach zero for values very far from the diagonal. The meaning of "very many" depends on Γ, the dimensionality of the coordinate space. If Ω is the total number of states, then $\omega = \Omega^{1/\Gamma}$ is the geometric mean of the number along each dimension. It is in the limit that this number, ω, becomes large that the density matrix approaches a product of a positive function $W(q')$ with the Dirac delta function product, $D(q' - q'')$. In section 7.10 we show that in this limit the average properties of the ensemble go over into those for a classical ensemble.

7.8. The Wigner representation

A somewhat peculiar transform, $W_Q(\boldsymbol{p},\boldsymbol{q})$, of the density matrix, first introduced by Wigner, gives a function of $\boldsymbol{p}$ and $\boldsymbol{q}$ which corresponds most closely to the classical probability density, $W_{\mathrm{Cl}}(\boldsymbol{p},\boldsymbol{q})$.

Before displaying this function some comments are necessary. The fundamental philosophy of quantum mechanics precludes the possibility of defining a probability density function in the combined $(\boldsymbol{p},\boldsymbol{q})$-space. The basis of quantum mechanics is that no equations can be given which answer questions that no conceivable experiment could verify. The uncertainty principle states that no conceivable experiment can give meaningful precise simultaneous values to a coordinate q_α *and* to its conjugate momentum, p_α. At best values within a range Δp_α and

Δp_α where the product $\Delta q_\alpha \Delta p_\alpha$ is of order h can have meaning. Since no experiment can ever fix the value of both p and q precisely there is no operational meaning to the question of what the probability is that a member system of an ensemble would have such precise $\boldsymbol{p},\boldsymbol{q}$-values. There is therefore no meaning to a probability density function, $W(\boldsymbol{p},\boldsymbol{q})$, defined at every point $\boldsymbol{p},\boldsymbol{q}$, in the Γ-space.

However, this does not preclude the possibility that a function $W_Q(\boldsymbol{p},\boldsymbol{q})$ might exist such that its integral

$$\iint_{\Delta p \Delta q} W_Q(\boldsymbol{p},\boldsymbol{q}) d\boldsymbol{p} d\boldsymbol{q}$$

over a range $\Delta\boldsymbol{p}\Delta\boldsymbol{q}$ for which every $\Delta p_\alpha \Delta q_\alpha > h$ might give the probability that a member system of the ensemble be found within this range. It is in this sense that the Wigner function $W_Q(\boldsymbol{p},\boldsymbol{q})$ "corresponds" to a probability density function.

It is therefore not surprising that the function $W_Q(\boldsymbol{p},\boldsymbol{q})$, which we shall define, has one property that no true probability density function could have. The function W_Q is indeed real (not complex), but it is not necessarily positive everywhere, and examples can be constructed for which it takes negative values at some points. No meaning could be given to a probability density which was negative.

The Wigner function is constructed as follows. Start with the coordinate representation $\Gamma(\boldsymbol{q}',\boldsymbol{q}'')$ of the density matrix. For each coordinate, α, transform q_α',q_α'' to q_α,q_α''', by the relations

$$q_\alpha = \tfrac{1}{2}(q_\alpha' + q_\alpha''), \qquad q_\alpha''' = q_\alpha' - q_\alpha'', \tag{7.8.1}$$
$$q_\alpha' = q_\alpha + \tfrac{1}{2}q_\alpha''', \qquad q_\alpha'' = q_\alpha - \tfrac{1}{2}q_\alpha''',$$

so that

$$dq_\alpha' dq_\alpha'' = dq_\alpha dq_\alpha'''. \tag{7.8.1'}$$

As before, use $\boldsymbol{q}$ to symbolize the set q_α and q''' for the set q_α'''. The Wigner function is

$$W_Q(\boldsymbol{p},\boldsymbol{q}) = h^{-\Gamma} \int \Gamma(\boldsymbol{q} + \tfrac{1}{2}\boldsymbol{q}''', \boldsymbol{q} - \tfrac{1}{2}\boldsymbol{q}''') e^{-(2\pi i/h)\boldsymbol{q}''' \cdot \boldsymbol{p}} d\boldsymbol{q}'''. \tag{7.8.2}$$

The inverse transform is then

$$\Gamma(\boldsymbol{q} + \tfrac{1}{2}\boldsymbol{q}''', \boldsymbol{q} - \tfrac{1}{2}\boldsymbol{q}''') = h^{-\Gamma} \int W_Q(\boldsymbol{p},\boldsymbol{q}''') e^{(2\pi i/h)\boldsymbol{q}''' \cdot \boldsymbol{p}} d\boldsymbol{p}. \tag{7.8.2'}$$

If, similarly to (1), we define for each α,

$$p_\alpha = \tfrac{1}{2}(p_\alpha' + p_\alpha''), \qquad p_\alpha''' = p_\alpha' - p_\alpha'', \tag{7.8.3}$$
$$p_\alpha' = p_\alpha + \tfrac{1}{2}p_\alpha''', \qquad p_\alpha'' = p_\alpha - \tfrac{1}{2}p_\alpha''',$$

$$dp_\alpha' dp_\alpha'' = dp_\alpha dp_\alpha''', \tag{7.8.3'}$$

and use p and p''' for the sets of momenta we have

$$q' \cdot p' - p'' \cdot q'' = (q + \tfrac{1}{2}q''') \cdot (p + \tfrac{1}{2}p''') - (q - \tfrac{1}{2}q'') \cdot (p - \tfrac{1}{2}p''')$$
$$= q \cdot p''' + p \cdot q''', \tag{7.8.3''}$$

so that the transform (7.6.13′) becomes

$$\Gamma(q + \tfrac{1}{2}q''', q - \tfrac{1}{2}q''')$$
$$= h^{-\Gamma} \int\int \Gamma(p' + \tfrac{1}{2}p''', p' - \tfrac{1}{2}p''')e^{(2\pi i/h)(p' \cdot q''' + q \cdot p''')}\, dp'dp''' \tag{7.8.4}$$

after relabeling $p \rightarrow p'$. Use this in (2), and integrate first over dq'''. From (7.6.8′) the integration gives $h^{\Gamma}D(p - p')$, so that subsequent integration over dp' simply sets $p' = p$ in $\Gamma(p' + \tfrac{1}{2}p''', p' - \tfrac{1}{2}p''')$. One finds

$$W_Q(p,q) = h^{-\Gamma} \int \Gamma(p + \tfrac{1}{2}p''', p - \tfrac{1}{2}p''')e^{(2\pi i/h)q \cdot p'''}\, dp''', \tag{7.8.5}$$

$$\Gamma(p + \tfrac{1}{2}p''', p - \tfrac{1}{2}p''') = h^{-\Gamma} \int W_Q(p,q)e^{-(2\pi i/h)q \cdot p'''} dq' \tag{7.8.5'}$$

Now let us examine the properties of this function W_Q. In the first place we see that W_Q is equal to its own conjugate complex, $W_Q{}^*$, since, from (2),

$$W_Q{}^* = h^{-\Gamma} \int \Gamma^*(q + \tfrac{1}{2}q''', q - \tfrac{1}{2}q''')e^{(2\pi i/h)q''' \cdot p}\, dq''' \tag{7.8.6}$$

and from the Hermitian character of $\mathbf{\Gamma}$ one has that

$$\Gamma^* = \Gamma(q - \tfrac{1}{2}q''', q + \tfrac{1}{2}q''').$$

Substituting $-q''' = q'$ one obtains

$$W_Q{}^* = h^{-\Gamma} \int \Gamma(q + \tfrac{1}{2}q', q - \tfrac{1}{2}q')e^{-(2\pi i/h)q' \cdot p}\, dq = W_Q, \tag{7.8.6'}$$

since the limits of integration of $q''' = -q'$ are even in q''', and this is also true when the coordinate range is intrinsically bounded. A function which is equal to its own conjugate complex is necessarily real.

From (2) and (7.6.8) we have

$$\int W_Q(p.q)dp = \int \Gamma(q - \tfrac{1}{2}q''', q - \tfrac{1}{2}q''')e^{-(2\pi i/h)q''' \cdot p}\, dq'''dp$$
$$= \int \Gamma(q + \tfrac{1}{2}q''', q - \tfrac{1}{2}q''')D(q''')dq = \Gamma(q,q), \tag{7.8.7}$$

and similarly from (5)

$$\int W_Q(p,q)dq = \Gamma(p,p). \tag{7.8.8}$$

We therefore have that the integral over p of $W_Q(q,p)$ gives the probability density, $\Gamma(q,q)$ in the coordinate space, and the integral over q of W_Q gives the probability density, $\Gamma(p,p)$ in the momentum space. It follows that for any function of $f(p)$ of the coordinates alone, or any function $g(q)$ of the momenta alone, the average values for the ensemble are given by the classical prescription using W_Q as a probability density,

$$\bar{f} = \iint f(q) W_Q,(p,q) dp dq, \tag{7.8.9}$$

$$\bar{g} = \iint g(p) W_Q(p,q) dp dq. \tag{7.8.10}$$

It will not be shown here, but it is not difficult to prove that for any mixed function $F(p,q)$ *which does not contain any product* $p_\alpha q_\alpha$ *of a single coordinate* q_α *with its* conjugated momentum, p_α, the prescription is the same, namely the average value $\bar{F}$ is the integral over p and q of the classical function F times W_Q. This, however, is not, and cannot be the case for the average value of any product $q_\alpha p_\alpha$, for the classical integral must necessarily give the same average for $q_\alpha p_\alpha$ as for $p_\alpha q_\alpha$. We have seen in section 7.5 that the quantum mechanical averages differ by $h/2\pi i$. It is not difficult to show that the integral of $q_\alpha p_\alpha$ times W_Q gives the arithmetical mean of the two quantum mechanical averages, and since $\overline{p_\alpha q_\alpha} - \overline{q_\alpha p_\alpha} = h/2\pi i$ it is easy to express either of the two quantum mechanical averages in terms of the classical integral, similar rules can be found for more complicated products of powers.

The fact that the quantum mechanical averages, $\overline{p_\alpha q_\alpha}$ and $\overline{q_\alpha p_\alpha}$ differ is itself a demonstration that no true probability density function, $W(p,q)$, can exist for the quantum mechanical ensemble. Were such a function to exist the two averages would necessarily be given by the classical integral, and would be identical. Nevertheless the Wigner function, $W_Q(p,q)$, "corresponds" closely to such a probability density, and in the classical limit, Planck's constant h approaching zero, does go over into the classical function.

7.9. The equilibrium density matrix

The density matrix has been discussed in the last four sections as a mathematical tool for describing any arbitrary ensemble of systems for which the quantum mechanical state, $\Phi_i(q)$, of each system i is assumed to be known. No assumption was made about whether or not the ensemble was one of systems at equilibrium. For an equilibrium ensemble of systems, each with a fixed number set N of molecules, at fixed volume V and temperature T the assertion has previously been made that the probability of observing a member system to be in the quantum state Ψ_K, which is a solution of the time independent Schrödinger equation,

$$\Psi_K(q) = E_K\Psi_K(q), \tag{7.9.1}$$

is

$$\Gamma_{K,K} = e^{(A-E_K)/kT}, \tag{7.9.2}$$

with A the Helmholtz free energy.

We need, here, to expand somewhat on the meaning of this assertion. We do *not* presume that if the state of a single member i of the ensemble were analysed it would necessarily be purely one of the solutions, Ψ_K, of (1). Rather, we assume that if its state Φ_i is expressed as a linear sum of these, as discussed at the end of section 7.5,

$$\Phi_i(q) = \sum_K a_{iK}\Psi_K(q), \tag{7.9.3}$$

then, in the limit of an infinite number, M, of systems,

$$\Gamma_{K,K} = \lim_{M\to\infty}[M^{-1}\sum_{i=1}^{i=M} a_{iK}a_{iK}{}^*] = e^{(A-E_K)/kT}. \tag{7.9.4}$$

In addition we make the "assumption of random phases," namely that if one writes the generally complex numbers a_{iK} as

$$a_{iK} = A_{iK}\exp i\phi_{iK}, \tag{7.9.5}$$

with A_{iK} real positive, and $0 < \phi_{iK} < 2\pi$, then the distribution of the "phase angles," ϕ_{iK}, is completely random, so that in the limit $M \to \infty$

$$\Gamma_{K',K''} = \lim_{M\to\infty}[M^{-1}\sum_i A_{iK'}A_{iK''}\exp i(\phi_{iK'} - \phi_{iK''})] = 0. \tag{7.9.6}$$

This means that the equilibrium ensemble has a density matrix which is diagonal in the representation K',K'' of the quantum numbers describing the solution Ψ_K to the Schrödinger equation (1). It is in this representation that the energy matrix is diagonal.

The great power of the density matrix formulation is that we can now evaluate, at least in principle, the elements in any other representation such as the coordinate or momentum representation without ever actually finding the solutions Ψ_K to (1). Define the operator expression $e^{-\mathscr{H}/kT}$, with $\mathscr{H}$ the Hamiltonian operator, to mean

$$e^{-\mathscr{H}/kT} = \sum_{n>0}\frac{(-)^n}{(kT)^n}\mathscr{H}^n \tag{7.9.7}$$

where the power, n, of the operator $\mathscr{H}$ simply means repetition of the operation n times. From (1) that $\mathscr{H}\Psi_K = E_K\Psi_K$ with E_K a number, it follows that

$$\mathscr{H}^n\Psi_K = E_K{}^n\Psi_K, \tag{7.9.7'}$$

$$e^{-\mathscr{H}/kT}\Psi_K(q) = e^{-E_K/kT}\Psi_K(q). \tag{7.9.8}$$

The Hamilton operator is real, so that $\Psi_K{}^*$, the conjugate complex of Ψ_K, is also a solution of (1) to the same energy E_K, and

$$e^{-\mathscr{H}/kT}\Psi_K{}^*(q) = e^{-E_K/kT}\Psi_K{}^*(q). \tag{7.9.8'}$$

Similarly, since

$$e^{-\mathscr{H}/2kT}\Psi_K(q) = e^{-E_K/2kT}\Psi_K(q), \tag{7.9.8''}$$

we have

$$\begin{aligned}[e^{-\mathscr{H}/2kT}\Psi_K(q')][e^{-\mathscr{H}/2kT}\Psi_K{}^*(q'')]& \\ = e^{-E_K/kT}\Psi_K(q')\Psi_K(q'').&\end{aligned} \tag{7.9.8'''}$$

Now the coordinate representation, $\mathbf{\Gamma}_q$ of the density matrix is obtained from the $\boldsymbol{K}$-quantum number representation, $\mathbf{\Gamma}_k$, by the unitary matrix Ψ of $\boldsymbol{q},\boldsymbol{K}$ elements $\Psi_K(\boldsymbol{q})$, and its conjugate transpose $\Psi\dagger$ of $\boldsymbol{K},\boldsymbol{q}$ elements $\Psi_K{}^*(\boldsymbol{q})$

$$\mathbf{\Gamma}_q = \boldsymbol{\Psi}\mathbf{\Gamma}_k\boldsymbol{\Psi}\dagger, \tag{7.9.9}$$

which, since $\mathbf{\Gamma}_k$ has only diagonal elements, means that

$$\Gamma(q',q'') = \sum_K \Psi_K(q')\Gamma_{K,K}\Psi_K{}^*(q''). \tag{7.9.9'}$$

Using (4) for $\Gamma_{K,K}$, and any of the equations (8), (8′) or, (8″), one may write

$$\Gamma(q',q'') = e^{(A-\mathscr{H})/kT}\sum_K \Psi_K(q')\Psi_K{}^*(q'') = e^{(A-\mathscr{H})/kT}\Psi\Psi\dagger, \tag{7.9.10}$$

where the Hamiltonian operator $\mathscr{H}$ can be understood to be either

$$\mathscr{H} = \mathscr{H}' = [(h/2\pi i)\nabla_{q'},q'], \tag{7.9.11}$$

operating *only* on $\boldsymbol{q}'$, or

$$\mathscr{H} = \mathscr{H}'' = \mathscr{H}[(h/2\pi i)\nabla_{q''},q''], \tag{7.9.11'}$$

operating only on $\boldsymbol{q}''$, or to mean

$$\mathscr{H} = \tfrac{1}{2}\mathscr{H}' + \tfrac{1}{2}\mathscr{H}'', \tag{7.9.11''}$$

as the sum of one-half of the operator on $\boldsymbol{q}'$ plus one-half of the operator on $\boldsymbol{q}''$. For some purposes the symmetric form (11″) is the most convenient.

Now we have already seen that for any complete orthonormal set of functions, $\Psi_K(q)$, the matrix $\mathbf{\Psi}$ is unitary,

$$\mathbf{\Psi}\mathbf{\Psi}\dagger = 1, \tag{7.9.12}$$

meaning that

$$\sum_K \Psi_K(q')\Psi_K(q'') = D(q' - q''), \tag{7.9.12$'$}$$

with D the multi-dimensional Dirac function in the coordinate system q. Hence (10) can be written as

$$\Gamma(q',q'') = e^{(A-\mathscr{H})/kT} D(q' - q''), \tag{7.9.13}$$

where the operator, $\mathscr{H}$, of (11), (11′) or (11″) operates on the unit matrix, 1, of diagonal elements $D(q' - q'')$ in the coordinate representation.

Now in order to express $D(q' - q'')$ of (13) in terms of analytical functions we may choose any other convenient complete orthonormal set of functions, $\Theta_M(q)$ as the q,M elements of a unitary matrix $\mathbf{\Theta}$, with $\Theta^*{}_M(q)$ as the M,q elements of its conjugate transpose, $\mathbf{\Theta}\dagger$,

$$\mathbf{\Theta}\mathbf{\Theta}\dagger = 1 \tag{7.9.14}$$

$$\sum_M \Theta_M(q')\Theta_M(q'') = D(q' - q'') \tag{7.9.14$'$}$$

to write (13) as

$$\Gamma(q',q'') = e^{(A-\mathscr{H}/kT)} \sum_M \Theta_M(q')\Theta^*{}_M(q'') = e^{(A-\mathscr{H}/kT)}\mathbf{\Theta}\mathbf{\Theta}\dagger. \tag{7.9.15}$$

The above argument may appear too facile, using the very singular product of Dirac functions as intermediary, and indeed it is, since it disguises one difficulty of maintaining proper symmetry in the functions. We will therefore prove (15) from (10) in a more conventional manner. The matrices

$$A = \mathbf{\Theta}\dagger\mathbf{\Psi}, \qquad A = \mathbf{\Psi}\dagger\mathbf{\Theta}, \tag{7.9.16}$$

are conjugate transposes of each other, since they have elements

$$a_{M,K} = \Theta^*{}_M \cdot \mathbf{\Psi}_K, \qquad a\dagger_{K,M} = \mathbf{\Psi}^*{}_K \cdot \Theta_M, \tag{7.9.16$'$}$$

respectively, and A is unitary since, from the unitary character of $\mathbf{\Psi}$ and $\mathbf{\Theta}$,

$$A\dagger A = \mathbf{\Psi}\dagger\mathbf{\Theta}\mathbf{\Theta}\dagger\mathbf{\Psi} = \mathbf{\Psi}\dagger\mathbf{1}\mathbf{\Psi} = \mathbf{\Psi}\dagger\mathbf{\Psi} = \mathbf{1}, \tag{7.9.17}$$

$$AA\dagger = \mathbf{\Theta}\dagger\mathbf{\Psi}\mathbf{\Psi}\dagger\mathbf{\Theta} = \mathbf{1}. \tag{7.9.17$'$}$$

In terms of the element summation these two equations are

$$\sum_M a\dagger_{K',M} a_{M,K''} = \delta(K' - K''), \tag{7.9.18}$$

$$\sum_M a_{M',K}, a_{K,M'} = \delta(M' - M''). \tag{7.9.18'}$$

From (16) we may write

$$\mathbf{\Theta} A = \mathbf{\Theta\Theta}\dagger\mathbf{\Psi} = \mathbf{1\Psi} = \mathbf{\Psi}, \tag{7.9.19a}$$

and, similarly,

$$\mathbf{\Psi} A\dagger = \mathbf{\Theta} \tag{7.9.19b}$$

$$A\dagger\mathbf{\Theta}\dagger = \mathbf{\Psi}\dagger \tag{7.9.19c}$$

$$A\mathbf{\Psi}\dagger = \mathbf{\Theta}\dagger \tag{7.9.19d}$$

which give, in matrix notation, the equations for developing one complete set of orthonormal functions in terms of another, namely,

$$\Psi_K(q) = \sum_M \Theta_M(q) a_{M,K}, \tag{7.9.19a'}$$

$$\Theta_M(q) = \sum_K \Psi_K(q) a\dagger_{K,M}, \tag{7.9.19b'}$$

$$\Psi_K{}^*(q) = \sum_M a\dagger_{K,M} \Theta^*{}_M(q), \tag{7.9.19c'}$$

$$\Theta^*{}_M(q) = \sum_K a_{K,M} \Psi_K{}^*(q). \tag{7.9.19d'}$$

Using (19a) and (19c) with (17′)

$$\mathbf{\Psi\Psi}\dagger = \mathbf{\Theta} AA\dagger\mathbf{\Theta}\dagger = \mathbf{\Theta 1 \Theta}\dagger = \mathbf{\Theta\Theta}\dagger, \tag{7.9.20}$$

which, in terms of the sums over the elements, is

$$\begin{aligned} &\sum_K \Psi_K(q')\Psi_K{}^*(q'') \\ &\quad = \sum_{M'}\sum_{M''} \Theta_{M'}(q')\Big(\sum_K a_{M',K} a\dagger_{K,M''}\Big)\Theta^*{}_{M''}(q'') \\ &\quad = \sum_M \Theta_M(q')\Theta^*{}_M(q'') \end{aligned} \tag{7.9.20'}$$

from (18′). Since the elements, $a_{M,K}$ and $a\dagger_{K,M}$ of the matrices A and $A\dagger$ are pure numbers they are not affected by the operator $\mathscr{H}$, and (20) can be used for $\mathbf{\Psi\Psi}\dagger$ behind the operator $e^{(A-\mathscr{H})/kT}$ of (10), leading to (15).

However, certain conditions are imposed on the solutions, Ψ_K, of the Schrödinger equation, and the members, Θ_M, of the complete orthonormal set, which from (19b′) contains only linear combinations of the Ψ_K's will necessarily be limited by these conditions. Most of these conditions are fairly trivial and easily satisfied, but one, the symmetry condition, causes difficulty.

The functions $\Theta_M(q)$ must be continuous, single valued, and differentiable in the space, q, allowed to the systems. They must form a complete orthonormal set in this space, which requires that they be square integrable. If cartesian coordinates are used the range minus to plus infinity is permissible, even for a bounded system of finite volume, but it is then necessary that the Hamiltonian operator $\mathscr{H}$ contain a potential term which becomes positive infinite if any molecule is outside the bounds of V. In practice this is awkward, and instead one uses functions which are orthogonal when integrated over the finite volume V. Strictly, boundary conditions that the functions are zero at the walls should be used, but as long as surface effects are not being investigated functions with "periodic boundary conditions" are adequate. An example is the functions $e^{2\pi inx/L}$ for integer n values when the limits on x are $0 \leqslant x \leqslant L$ or $-\frac{1}{2}L \leqslant x \leqslant \frac{1}{2}L$. All these conditions are easy enough to satisfy.

In addition, the allowed solutions to the Schrödinger equation (1) must be symmetric in permutations of the coordinates of identical Bose-molecules, and antisymmetric in the exchange of Fermi-molecule pairs. Now if the coordinate system q' is written in the conventional manner, $q' = q'_1, q'_2, \ldots, q'_\alpha, \ldots, q_N'$, with q'_α the coordinates of the molecule α, so that the coordinate q'' with $q_1'' = q_2', q_2'' = q_1', \ldots, q_\alpha' = q_\alpha'', \ldots$, is different from q' then the purely symmetric functions Ψ_K do *not* form a complete set in this coordinate space, since only completely symmetric functions can be expressed as a linear combination of these. As a result the sum (12′) is no longer the function $D(q' - q'')$ defined by the product (7.5.21) of all Dirac delta functions. A coordinate system, Q_s (for which there is no simple convenient notation when the coordinates are continuous, and the individual molecular coordinates q_α are more than one-dimensional) *does* exist for which the completely symmetric functions form a complete set. This is the coordinate system $Q_s = q_1, \ldots, q_\alpha, \ldots q_N$, in which, however, we regard the order of writing the subscripts as immaterial, that is in which we regard all $N!$ permutations of the coordinate positions as identical values of Q_s. The sum (12′) of purely symmetric functions Ψ_K is a function, $D(Q_s' - Q_s'')$, which is non-zero only when $Q_s'' \equiv Q_s'$, and for which the integral over dQ_s'' is unity. However, we must interpret $Q_s \equiv Q_s''$ for every one of the $N!$ permutations for which $q_\alpha' \equiv q_\beta''$ with some

$\beta(\alpha)$ for every one of the N values of α. This function, $D(\boldsymbol{Q}_s' - \boldsymbol{Q}_s'')$ can be expressed as a function, $D_s(\boldsymbol{q}' - \boldsymbol{q}'')$ of the more conventional coordinates, $\boldsymbol{q}$, and is constructed from the product function, D, defined by (7.5.21) by taking the average of its value for all $N!$ permutations of either $\boldsymbol{q}'$ or $\boldsymbol{q}''$,

$$D_s(\boldsymbol{q}' - \boldsymbol{q}'') = \frac{1}{N!} \sum_{p=1}^{p=N!} \mathscr{P}_p' D(\boldsymbol{q}' - \boldsymbol{q}'') \tag{7.9.21}$$

where $\mathscr{P}_p'$ is the permutation operator that permutes the coordinates of the individual molecules in $\boldsymbol{q}'$.

The coordinate system $\boldsymbol{Q}_s$ is the natural one for describing a configuration of *identical* molecules, the position $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$ not meaning that molecule number 1 is at $\boldsymbol{q}_1$ and number 2 at $\boldsymbol{q}_2, \ldots$, but that a molecule is at $\boldsymbol{q}_1$, another at $\boldsymbol{q}_2$, etc. For three-dimensional cartesian coordinates limited to a volume V, and N molecules, the volume of integration within V of $d\boldsymbol{Q}_s$ is then $V^N/N!$ instead of V^N for the integral of $d\boldsymbol{Q}$. If the coordinates of the single molecules are one-dimensional cartesian coordinates limited to the range $-\frac{1}{2}L \leqslant x_\alpha \leqslant \frac{1}{2}L$, then one can use the conventional numbered coordinates, $x_1, x_2, \ldots, x_\alpha, \ldots, x_N$, and the limitation $-\frac{1}{2}L \leqslant x_1 \leqslant x_2 \leqslant \ldots \leqslant x_\alpha \leqslant \ldots \leqslant x_N < \frac{1}{2}L$, meaning that x_1 is the coordinate of the molecule that has the lowest coordinate value, x_2 that of the next, etc. However, even this method of designation of $\boldsymbol{Q}_s$ would be awkward, since a complete set of orthonormal functions is not easy to find in a space with such peculiar boundary conditions. In the analogous problem of writing the discrete quantum number $\boldsymbol{K}$ describing the symmetric state $\Psi_{\boldsymbol{K}}$, we found a notation in eq. (1.5.5) by specifying $n(\boldsymbol{K})$, the number of times a molecular quantum number $\boldsymbol{K}$ appeared in the set $\boldsymbol{K}_1, \boldsymbol{K}_2, \ldots, \boldsymbol{K}_\alpha, \ldots, \boldsymbol{K}_N$, thus obviating the necessity of distinguishing the subscripts on a given $\boldsymbol{K}$. A method which resembles this, and can, in principle, be used to approach a continuous coordinate system is to divide the volume V into a lattice of contiguous cells, numbered by discrete values of a vector $\boldsymbol{K}$, and to specify the number $n(\boldsymbol{K})$ of molecules in each cell. As the cells approach zero volume, and become infinite in number, $n(\boldsymbol{K}) = 0$ or 1 only, the description approaches that of the coordinate system $\boldsymbol{Q}_s$. In practice this method, the cell method, has only been used for cells of the order of the molecular volume in size.

We therefore resign ourselves to use the more convenient, but not entirely natural, coordinate system $\boldsymbol{q}$, with numbered $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$, etc.

As a result, in equations involving integration over dq the factor $N!$ always shows up in the denominator in all final equations.

In practice the "convenient" orthonormal function set $\Theta_M(q)$ is chosen as a product of functions, $\theta_{m\alpha}(q_\alpha)$ of the coordinates of molecule α, and indeed $\theta_{m\alpha}$ is taken as a product of functions of the single coordinates of the molecule α. The quantum number M of Θ_M is then taken as

$$M = m_1, m_2, \ldots, m_\alpha, \ldots, m_N, \tag{7.9.22}$$

and the sum over all M values in (15) is taken as the independent sum over all m_α's. This, then, includes the states of all symmetry character. The rough rule that only the fraction $1/N!$ of them will have the proper symmetry is then used to divide the sum by $N!$, obtaining an "unsymmetrized" result,

$$\Gamma_u(q',q'') = \frac{1}{N!}\, e^{(A-\mathscr{H})/kT} \prod_{\alpha=1}^{\alpha=N} \sum_{m_\alpha} \theta_{m\alpha}(q_\alpha{}')\theta_{m\alpha}{}^*(q_\alpha{}''). \tag{7.9.23}$$

This function, $\Gamma_u(q',q'')$, is the coordinate representation of the equilibrium "Boltzmann" density matrix. Setting its trace equal to unity gives the Boltzmann value for the Helmholtz free energy A.

To correct Γ_u for Bose–Einstein or Fermi–Dirac systems we can now "symmetrize" and find the correct function. Any function of q can be written as a sum of functions, one of which is totally symmetric in the exchange of all pairs q_α, q_β, one of which is totally antisymmetric in all pair exchanges, and the others are antisymmetric in the pair exchanges of a subset $\{n\}_N$ with $n \leqslant 2$, and have no special symmetry in exchanges of pairs of the remaining $N - n$ molecular coordinates. If the sum of the function for all $N!$ permutations of the coordinates is taken then those component functions which are antisymmetric in any pair add to zero, since equal numbers of plus and minus terms appear differing only in the exchange of that pair. The sum is then $N!$ times the symmetric component. Thus the symmetric component, $f_s(q)$, is

$$f_s(q) = \frac{1}{N!} \sum_{p=1}^{p=N!} \mathscr{P}_p f(q) \tag{7.9.24}$$

where $\mathscr{P}_p$ is the operator that permutes the coordinates. Similarly one can show that the totally antisymmetric component is

$$f_a(q) = \frac{1}{N!} \sum_{p=1}^{p=N!} (-)^p \mathscr{P}_p f(q), \tag{7.9.24'}$$

where p, the index of the permutation, is odd if an odd number of pairs are exchanged in the permutation, and even for an even number of exchanges.

Now the function $\Gamma_u(q',q'')$ of (23) is necessarily symmetric in any permutation of *both* q' and q'' since the Hamiltonian operator is necessarily identical for identical molecules, and hence symmetric in the exchange of their positions. A permutation $\mathscr{P}_p$ of q', followed by another permutation $\mathscr{P}_{p''}$ of q'' can alway be obtained by first making the permutation $\mathscr{P}_{p''}$ on *both* q' and q'', and then some other permutation $\mathscr{P}_{p'}$ on q' alone. The sums are also equal, namely

$$\left(\frac{1}{N!}\right)^2 \sum_p \mathscr{P}_p(q') \sum_{p''} \mathscr{P}_{p''}(q'')$$

$$= \left[\frac{1}{N!} \sum_p \mathscr{P}_{p''}(q',q'')\right]\left[\frac{1}{N!} \sum_{p'} \mathscr{P}_{p'}(q')\right]. \qquad (7.9.25)$$

Since the permutation of both q' and q'' with the same permutation leaves $\Gamma_u(q',q'')$ unchanged, the first operation leads to $\mathbf{\Gamma}_u$, and $1/N!$ times the sum of permutations on only one of the coordinates, q', (or only on q'') is then the same as selecting the component function of $\Gamma_u(q',q'')$ which is totally symmetric in both q' and q''. Since $\mathbf{\Gamma}_u$ in (23) has already been arbitrarily divided by $N!$ one has for the coordinate representation of the density matrix for an equilibrium ensemble of Bose–Einstein molecules,

$$\Gamma_{\text{B.E.}} = \sum_{p=1}^{p=N!} \mathscr{P}_p{}'\Gamma_u(q',q''), \qquad (7.9.26)$$

where the permutation operator acts on q' alone. Similarly for an ensemble of Fermi–Dirac molecules at equilibrium

$$\Gamma_{\text{F.D.}} = \sum_{p=1}^{p=N!} (-)^p \mathscr{P}_p{}'\Gamma_u(q',q''). \qquad (7.9.26')$$

7.10. The classical limit

In this section we shall show that in the limit of setting Planck's constant, h, equal to zero, the condition that the trace of the density matrix, $\mathbf{\Gamma}$, is equal to unity leads to the classical equations for an ensemble of fixed V, N, T. We also find that in this limit the coordinate representation $\Gamma(q',q'')$ is diagonal and differs from zero only at $q' = q''$.

Neglect possible spin coordinates and use cartesian coordinates $x_{\alpha i}$ with $1 \leqslant \alpha \leqslant N$, and $i = 1, 2, 3$, for the coordinate of the atoms of the systems. If the systems are in cubic boxes of edge length L, $V = L^3$, we can limit the coordinate range to $-\frac{1}{2}L \leqslant x_{\alpha i} \leqslant \frac{1}{2}L$ and use a product of plane wave functions,

$$\phi_{k\alpha i}(x_{\alpha i}) = (1/\sqrt{L}) \exp (2\pi i k_{\alpha i} x_{\alpha i}/L), \tag{7.10.1}$$

$$\Theta_M(\boldsymbol{q}) = \prod_{\alpha=1}^{\alpha=N} \prod_{i=1}^{i=3} \phi k_{\alpha i}(x_{\alpha i}), \tag{7.10.1$'$}$$

$$\boldsymbol{M} = k_{11}, k_{12}, \ldots, k_{\alpha i}, \ldots, k_{N3}. \tag{7.10.2}$$

As before, introduce the momentum $p_{\alpha i} = hk_{\alpha i}/L$, and substitute integration over $p_{\alpha i}$ for summation over $k_{\alpha i}$,

$$k \sum_{\alpha i} = (L/h) \int_{-\infty}^{+\infty} dp_{\alpha i}. \tag{7.10.2$'$}$$

Using the notation $\boldsymbol{p}$ for the set of momenta, and

$$\boldsymbol{p} \cdot \boldsymbol{q} = \sum_{\alpha=1}^{\alpha=N} \sum_{i=1}^{i=3} p_{\alpha i} \cdot x_{\alpha 1}, \tag{7.10.3}$$

$$\Theta_M(\boldsymbol{q}')\Theta_M(\boldsymbol{q}'') = \frac{1}{V^N} \Phi[\boldsymbol{p} \cdot (\boldsymbol{q}'' - \boldsymbol{q}')] = \frac{1}{V^N} e^{(2\pi i/h)\boldsymbol{p} \cdot (\boldsymbol{q}'' - \boldsymbol{q}')}, \tag{7.10.3$'$}$$

we may write (7.9.23) as

$$\Gamma(\boldsymbol{q}', \boldsymbol{q}'') = \frac{1}{N! h^{3N}} \int e^{(A - \mathscr{H})/kT} \Phi[\boldsymbol{p} \cdot (\boldsymbol{q}'' - \boldsymbol{q}')] d\boldsymbol{p}. \tag{7.10.4}$$

By definition we mean by the exponent of the operator,

$$e^{-\mathscr{H}/kT} = \sum_{n > 0} (-1/kT)^n \mathscr{H}^n, \tag{7.10.5}$$

where $\mathscr{H}^n$ means n-fold operation by the Hamiltonian operator. We use the symmetric form (7.9.11″) for $\mathscr{H}$ as one-half of the sum of the operator on q'' plus that on q', so that, for the cartesian coordinates

$$\mathscr{H} = \left\{ \sum_{\alpha=1}^{\alpha=N} \sum_{i=1}^{i=3} \frac{1}{4m_\alpha} \left(\frac{h}{2\pi i} \right)^2 \left[\left(\frac{\partial}{\partial x_{\alpha i}''} \right)^2 + \left(\frac{\partial}{\partial x_{\alpha i}'} \right)^2 \right] + \tfrac{1}{2}[U(q'') + U(q')] \right\}. \tag{7.10.6}$$

The first term of the sum (5), having $n = 0$, means no operation, and

$$(\mathscr{H})^0 \Phi = \Phi. \tag{7.10.7}$$

The next term, for which $n = 1$ produces

$$\mathscr{H}\Phi = H\Phi \tag{7.10.8}$$

where H is the classical Hamiltonian,

$$H = \left\{ \sum_{\alpha=1}^{\alpha=N} \sum_{i=1}^{i=3} \frac{p_{\alpha i}^{2}}{2m_\alpha} + \frac{1}{2}[U(q'') + U(q')] \right\}, \tag{7.10.9}$$

in which, however, the potential energy is the mean of the two values at q'' and q'. Now operate again with $\mathscr{H}$ on the right hand side of (8) and we find

$$\mathscr{H}^2\Phi = \left\{ H^2 + \sum_{\alpha=1}^{\alpha=N} \sum_{i=1}^{i=3} \left[\frac{1}{4m_\alpha} p_{\alpha i} \left(\frac{h}{2\pi i} \right) \left(\frac{\partial U(q'')}{\partial x_{\alpha i}''} - \frac{\partial U(q')}{\partial x_{\alpha i}'} \right) \right.\right.$$

$$\left.\left. - \frac{h^2}{16\pi^2 m} \frac{1}{2} \left(\frac{\partial^2 U(q'')}{\partial x_{\alpha i}''^2} + \frac{\partial^2 U(q')}{\partial x_{\alpha i}'^2} \right) \right] \right\} \Phi. \tag{7.10.10}$$

In the limit $h \to 0$ the terms involving derivatives of the potential energy disappear. Similarly for subsequent operations by $\mathscr{H}$,

$$\mathscr{H}^n\Phi = [H^n + O(h)]\Phi, \tag{7.10.11}$$

$$\lim_{h \to 0} [\mathscr{H}^n\Phi] = H^n\Phi, \tag{7.10.11$'$}$$

and so

$$\lim_{h \to 0} [e^{-\mathscr{H}/kT}\Phi] = e^{H/kT}\Phi, \tag{7.10.12}$$

where the Hamiltonian H is the classical symmetric Hamiltonian function (9) of the momenta $\boldsymbol{p}$ and two coordinate values, $\boldsymbol{q}'$, $\boldsymbol{q}''$ in the matrix. We now have, from (4),

$$\lim_{h \to 0} [\Gamma(\boldsymbol{q}',\boldsymbol{q}'')] = \frac{1}{N!} e^{[A - \frac{1}{2}U(q') - \frac{1}{2}U(q'')]/kT} J(\boldsymbol{q}',\boldsymbol{q}'') \tag{7.10.13}$$

with

$$J = \prod_{a=1}^{\alpha=N} \prod_{i=1}^{i=3} I(x_{\alpha i}', x_{i\alpha}'') \tag{7.10.14}$$

$$I(x', x'') = h^{-1} \int_{-\infty}^{\infty} dp \, e^{-(p^2/2mkT) + (2\pi i/h)p(x' - x'')}. \tag{7.10.15}$$

Now introduce the de Broglie wavelength, $\lambda_\alpha = h(2\pi m_\alpha kT)^{-1/2}$, which depends on the mass, m_α, of atom α. With this, and a new dummy variable

$$z = (\lambda p/h) - [i(x'' - x')/\lambda], \tag{7.10.16}$$

we can write the exponent in the integrand of (15) as

$$\begin{aligned}(p^2/2mkT) &+ (2\pi i/h)p(x'' - x') \\ &= -\pi\{\lambda^2 p^2/h^2 - 2(\lambda p/h)[i(x'' - x')/\lambda]\} \\ &= -\pi z^2 - \pi[(x'' - x')/\lambda]^2. \end{aligned} \tag{7.10.17}$$

With $h^{-1}dp = \lambda^{-1}dz$, and integration of z along the line $-i[(x'' - x')/\lambda]$ from minus to plus infinity, $\int e^{-\pi z^2}dz = 1$, we find that

$$I_{i\alpha} = \lambda_\alpha^{-1} \exp - \pi(x_{i\alpha}' - x_{i\alpha}'')^2/\lambda^2. \tag{7.10.18}$$

In the limit h decreases towards zero, λ_α approaches zero, and each factor $I_{i\alpha}$ becomes zero except when $x_{i\alpha}' = x_{i\alpha}''$. We then have that in the limit $h \to 0$ the coordinate representation, $\Gamma(\boldsymbol{q}',\boldsymbol{q}'')$, of the density matrix has zero off diagonal elements. Along the diagonal, $\boldsymbol{q}' = \boldsymbol{q}''$ it has the value

$$\lim_{h\to 0} [\Gamma(q,q)] = [\prod_\alpha N_\alpha! \lambda_\alpha^{3N_\alpha}]^{-1} \exp [A - U(q)]/kT. \tag{7.10.19}$$

With the configuration integral

$$Q_\tau = \frac{1}{N!}\int dq\, e^{-U(q)/kT} \tag{7.10.20}$$

of (2.4.4) we find from the normalization

$$\int \Gamma(q,q)dq = 1 \tag{7.10.21}$$

that

$$A = \sum_a 3N_a kT \ln \lambda_a - kT \ln Q_\tau \tag{7.10.22}$$

in agreement with the classical formulation of (2.4.5′).

Now, of course, strictly, since we have assumed $h \to 0$ and $\lambda_a \to 0$ we have an infinite expression for the Helmholtz free energy A. Forming $S = -(\partial A/\partial T)_{V,N}$ and $E = A + TS$ one finds that, since $\lambda \sim T^{-1/2}$, $E = \sum_a \frac{3}{2}N_a kT + kT(\partial \ln Q_\tau/\partial \ln T)_{V,N}$ which is finite. The entropy is infinite.

The reason for the infinite entropy is simple. With $h \to 0$ and the number of quantum states in any range going as $\Delta p_{\alpha i}\Delta x_{\alpha i}/h$ the number of states in any fixed classical portion of phase space increases without limit. The entropy is $S = k \ln \Omega$ with Ω the number of quantum states so that S goes logarithmically to infinity as h approaches zero, and with the factor $k\Gamma$ equal to k times the number of degrees of freedom, which is $3Nk$ in this case. The expression $S + k\Gamma \ln h$ remains finite.

Finally we also want to point out that for molecules having a finite range of repulsion the symmetrization adds only zero terms in the classical limit. The density matrix elements $\Gamma(\boldsymbol{q}',\boldsymbol{q}'')$ are zero unless $\boldsymbol{q}' = \boldsymbol{q}''$ which means that every molecule must be at the same position in $\boldsymbol{q}'$ as in $\boldsymbol{q}''$. But if two molecules are exchanged in $\boldsymbol{q}'$ then $\boldsymbol{q}' \neq \boldsymbol{q}''$ unless they were at identical positions originally. The large repulsive pair energies makes $\Gamma(\boldsymbol{q},\boldsymbol{q})$ zero when any two molecules are at the same position. Hence in (7.9.26) or (7.9.26′) only the first term, $p = 0$ with the identical permutation that permutes nothing is non-zero. All systems become "Boltzmann."

7.11. Perfect Bose gas cluster function

In section 7.2 it was shown that the equation of state of the perfect Bose gas could be written as

$$P/kT = \sum_{n \geqslant 1} b_n z^n, \tag{7.11.1}$$

with

$$b_n = n^{-5/2}\lambda^{(3-3n)}, \tag{7.11.2}$$

eqs. (7.2.9), (7.2.10) and (7.2.14). It was promised in that section that we should show that the analogy to the cluster integral equation was not a pure numerical coincidence. We shall proceed to do this using the method of the last section, which is extremely simple for real h-values and the perfect gas case that the potential energy is identically zero.

In the last section the normalized unsymmetrized density matrix element, $\Gamma(\boldsymbol{q}',\boldsymbol{q}'')$ for an ensemble of systems of fixed N, V, T was given by (7.10.13) for the limit $h \to 0$. The trace of this matrix is unity, and this determined the Helmholz free energy, A. If we set $U \equiv 0$ the expression (7.10.13) is exact for all h-values, but the symmetrization sum of eq. (7.9.26) must be applied before the trace is taken. Let us first transform the expression of (7.10.13) to make it correspond as closely as possible to the form which was used in Chapter 4 for the imperfect gas.

From (1.4.12) that $A(V,T,N) = -kT \ln Q_c(V,T,N)$ and (4.1.1) with (4.1.2) relating Q_c to Q_τ we have

$$e^{-A/kT} = Q_c(N,V,T) = \lambda^{-3N} Q_\tau / N\,!, \tag{7.11.3}$$

$$Q_\tau = \iint_V \cdots \int e^{-U/kT} d\{N\}. \tag{7.11.4}$$

Comparing this with (7.10.13) and the condition that the trace of Γ is unity we see that, for the symmetrized perfect Bose gas,

$$Q_\tau = \{ \sum_{p=1}^{p=N\,!} \mathscr{P}_p\, \lambda^{3N} J(\boldsymbol{q}',\boldsymbol{q}'') \}_{\text{tr}}, \tag{7.11.5}$$

with

$$\lambda^{3N} J(\boldsymbol{q}',\boldsymbol{q}'') = \exp - \pi \sum_{i=1}^{i=N} [|\boldsymbol{r}_i' - \boldsymbol{r}_i''|/\lambda]^2. \tag{7.11.6}$$

The first unpermuted term $p = 0$ in the sum is unity on the diagonal $\boldsymbol{q}' \equiv \boldsymbol{q}''$. In a term in which the atom i has replaced j in the permutation we will have a factor $\exp - \pi[|\boldsymbol{r}_i' - \boldsymbol{r}j''|/\lambda]^2$ in $\mathscr{P}_p \lambda^{3N} J(\boldsymbol{q}',\boldsymbol{q}'')$ which, when $\boldsymbol{q}' = \boldsymbol{q}'' = q$, becomes

$$f(r_{ij}) = \exp - \pi[|\boldsymbol{r}_i - \boldsymbol{r}_j|/\lambda]^2. \tag{7.11.7}$$

Every permutation of the N molecules can be analysed as a product of cycle permutations. A cycle permutation of n numbered objects is one in which, say, object 1 replaces 2, object 2 replaces 3, . . ., object $n - 1$ replaces n, and n replaces 1. The factor introduced into the diagonal part, $\boldsymbol{q}' = \boldsymbol{q}'' = \boldsymbol{q}$, of $\mathscr{P}_p \lambda^{3N} J$ by such a single cycle is

$$C_n(r_1, r_2, \ldots, r_n) = f(r_{12}) f(r_{23}) \ldots f(r_{n-1,n}) f(r_{n,1}), \tag{7.11.8}$$

which is a function that, for finite n, approaches zero unless all n molecules are close together. We define a "cluster" function g_n as the sum of the $(n - 1)\,!$ different ways of cycling the same n molecules

$$g_n = \sum_{\alpha=1}^{\alpha=(n-1)\,!} C_{n\alpha}(r_1, \ldots). \tag{7.11.9}$$

One now has, from (5), with $g_1 \equiv 1$, that

$$Q_\tau(N,V,T) = \iint_V \cdots \int (\Sigma\{k\{n_i\}_N\}_u)[\prod_{i=1} g_{ni}\{n_i\}_N] d\{N\}. \tag{7.11.10}$$

As in section 4.2 we define the cluster integral by

$$b_n = \frac{1}{Vn!}\iint_V \cdots \int g_n\{n\}d\{n\}$$

$$= \frac{1}{n!}\iint \cdots \int [f(r_{12}) \ldots f(r_{n1})]dr_2 \ldots dr_n. \quad (7.11.11)$$

The rest of the procedure of (4.2), using (9) in the Grand Canonical partition function equation, (4.1.6) follows and as in eq. (4.2.12) we have

$$P/kT = \sum_{n\geqslant 1} b_n z^n. \quad (7.11.12)$$

The integral b_n of (11) can be evaluated by folding as in section 4.8, eqs. (4.8.10) to (4.8.14) with f given, now, by (7). One has

$$g(t) = \int 4\pi r^2 \, e^{-\pi(r/\lambda)^2} \frac{\sin 2\pi tr}{2\pi tr} \, dr$$

$$= \lambda^{-3} \, e^{-\pi\lambda^2 t^2}, \quad (7.11.13)$$

$$b_n = \frac{1}{n} \int 4\pi t^2 [g(t)]^n dt$$

$$= n^{-5/2} \lambda^{(3-3n)} \quad (7.11.14)$$

in agreement with (2).

One sees then that there is actually a statistical attraction in the coordinate space for the Bose gas molecules into clusters, and the classical treatment can be used, with the somewhat peculiar cluster function of eqs. (8), (9).

CHAPTER 8

DERIVATION OF THE PARTITION FUNCTION EQUATIONS

8.1. Outline of the derivation

We shall here undertake the derivation of the relations between the partition functions and the thermodynamic quantities. In so doing the laws of thermodynamics will appear as a consequence of the laws of mechanics, and not as a separate necessary axiomatic assumption of physical science. The derivation follows in successive steps according to the following outline.

1. The possibility of a thermodynamic fluctuation, not anticipated in the classical thermodynamic laws, is recognized, and the ensembles, consisting of the totality of all systems prepared according to specified fixed sets of thermodynamic variables are introduced. The average properties of the ensemble are then to be computed, not the certain property of a single system (section 2).
2. The ensemble corresponding to fixed volume, energy, and numbers of molecules of different species, called the microcanonical, is shown to be that of equal probability of all quantum states of that energy (sections 3, 4).
3. The quantity S, equal to k times the logarithm of the number of states available to the microcanonical ensemble is defined (section 5).
4. The quantity S is shown to be extensive, and to increase in all natural changes permitted to the ensemble at constant V, E, N (section 5).
5. The quantities $T = (\partial S/\partial E)_{V,N}{}^{-1}$ and $P = T(\partial S/\partial V)_{E,N}$ are shown to be temperature and pressure, respectively, which identifies S with the thermodynamic entropy (section 6).
6. The assumption that only conservative forces between molecules exist in nature requires the first law of thermodynamics. Statement (4) above, with (5) gives the second law. The fact that S is proportional to the logarithm of an integer, whose lowest possible value is unity, fixes a natural zero of S, which is the third law. The laws of thermodynamics are established, (section 7).

7. By considering an ensemble of supersystems, each of fixed $V, E, \boldsymbol{N}$, each of which consists of a system of large (infinite) size, in equilibrium with its subsystems, we compose ensembles of systems in equilibrium with a reservoir. The energies of the systems may fluctuate, the average values being determined by the temperature, $(\partial S/\partial E)^{-1}$, of the reservoir. One may also include cases where the walls of the systems are permeable to molecules of some or all species, so that the number set $\boldsymbol{N}$ in the systems may fluctuate. The properties are then determined by the set $\boldsymbol{\mu}$ of chemical potentials in the reservoir. A variety of ensembles may be composed, including the macrocanonical with fixed V, T, $\boldsymbol{N}$, and the grand canonical of fixed V, T, $\boldsymbol{\mu}$, as well as one of fixed P, T, $\boldsymbol{N}$. From each of these we derive an equation giving the appropriate thermodynamic potential for these variables as $-kT$ times the logarithm of a partition function.

8.2. The ensemble concept

The thermodynamic state of a system is given by specifying a very few variables. For instance the extensive variables, volume, V, energy, E, and the number set $\boldsymbol{N} = N_a, N_b, \ldots$, of molecules of species a, $b, \ldots$, respectively, are sufficient to describe the thermodynamic state at complete equilibrium. Now consistent with this thermodynamic state the system may be in any of a tremendously large number, Ω, of single quantum states, $\boldsymbol{K}$. In general the specification of $\boldsymbol{K}$ requires the specification of as many components as the number, Γ, of degrees of freedom in the system, a number of the order 10^{20} or greater.

The single quantum states $\boldsymbol{K}$ include those having rather exceptional average properties. For instance in a gas they include states in which all the molecules move along the z-axis only, and hence exert a positive pressure on the walls normal to this axis alone. True, that such a state would be expected to have a very evanescent time duration, collisions would be expected to produce random directions of motion in a very short time, but we cannot exclude the possibility that a single pressure measurement would give the "wrong" value. Nor are all possible abnormal states such that they would relax in very short time durations. For example, consider a system consisting of two equal subvolumes, $V_1 = V_2 = \frac{1}{2}V$, connected by a small opening, in which the total numbers, $N_a = N_b$, of components a and b in equal amounts were fixed. Some quantum states correspond to a large excess of one component in one subvolume, and a corresponding deficit in the other. This fluctuation would persist for a considerable time.

We therefore deny the classical thermodynamic concept that fixing intensive variables as P, T, or the set $\boldsymbol{\mu}$ of chemical potentials, μ_a, μ_b, . . ., should *precisely* determine such conjugate extensive variables such as V, E, or the number set $\mathbf{N}$ of molecules, and admit the possibility of fluctuations. We forego the pleasure of attempting to predict the properties of a single system, but instead introduce an ensemble, for which we attempt to compute the *average* properties of the members. In so doing we find that we can also predict the probability that an experimentally detectable deviation from the average property will be observed, and in accordance with common experience we find this probability, normally, to be vanishingly small.

The ensemble, then, consists of a large number, M, of systems, and in the limit we let M approach infinite value. The ensemble is intended to represent the totality of all systems prepared in the laboratory according to prescribed macroscopic directions. We may, for instance, suppose that it represents all systems for which V, E, and $\mathbf{N}$ are fixed, and no other peculiar conditions that are not explicitly specified are imposed. As example, in the case just mentioned let $V = V_1 + V_2$, and (1) and (2) be two flasks, connected by a small opening, and let $\mathbf{N} = N_a, N_b$ and $N_a = N_b$. If the experimental conditions are that the system is prepared with the N_a molecules of a in V_1, and the N_b of b in V_2, this must be defined in the ensemble. Since we limit ourselves to macroscopically stationary ensembles we then also impose the condition that a stopcock closes the connection between the two flasks in such a case. However, just as in thermodynamic considerations, we shall allow the stopcock to be opened, permitting a process to occur, but will limit our considerations to the final result of such a process. We do not discuss the time required for equilibrium to establish itself. In these respects our ensembles are described in exactly the same way that thermodynamic† systems are described, and with the same limitations to equilibrium, or to quasi-equilibrium. All processes are regarded as either infinitely slow, or as having proceeded to equilibrium.

However, in one respect the description of the possible ensembles differs from thermodynamic concepts. Since classical thermodynamics presupposes that if $V, E, \mathbf{N}$ are given, the intensive variables P, T, $\boldsymbol{\mu}$ are determined, the *state* is the same whether V, E, $\mathbf{N}$ are prescribed, or P, E, $\mathbf{N}$, or V, T, $\mathbf{N}$, or V, T, $\boldsymbol{\mu}$, provided consistent values are chosen.

---

† We use thermodynamic here in the classical sense, for which the word thermostatic might be more appropriate.

The statistical ensembles are *not* the same in these different cases, although their average properties will be seen to be the same.

For instance the thermodynamic system prescribed by fixed V, T, N corresponds to a system of fixed V, N, prepared by bringing it in equilibrium with a large heat reservoir whose energy is such that $(\partial S/\partial E)^{-1} = T$ has the prescribed value. In such a combined system, heat reservoirs plus the system in which we are interested, the allowed quantum states include those in which abnormally high, or abnormally low energies are in the interesting subsystem. The ensemble of fixed V, T, N, then, includes member systems of various energies.

The ensemble for which V, E, N are fixed is called the microcanonical, that for which V, T, N are fixed is known as the petit canonical, or simply canonical, and that of fixed V, U, μ is the grand canonical. Other ensembles are definable, particularly that of fixed P, T, N, but the first three are those most commonly used.

8.3. The microcanonical ensemble, the Liouville theorem

The microcanonical ensemble represents the totality of all thermodynamic systems for which V, N are fixed and the energy lies between E and $E + \Delta E$. The ΔE is the experimental uncertainty in the energy. We shall now demonstrate that this ensemble is one with equal probability of all quantum states, $\boldsymbol{K}$, with energies E_K, for which $E \leqslant E_K \leqslant E + \Delta E$. The corresponding classical statement is that the probability density, $W(\boldsymbol{p}^{(N)}, \boldsymbol{q}^{(N)})$, in the phase space, $\boldsymbol{p}^{(N)}$, $\boldsymbol{q}^{(N)}$, of the momenta and coordinates of the system, is a constant within the region for which $E \leqslant H(\boldsymbol{p}^{(N)}, \boldsymbol{q}^{(N)}) \leqslant E + \Delta E$.

The demonstration, subject to a certain interpretation of the ensemble which we discuss in the next section, is almost trivial for the quantum mechanical statement. The transition probability, $A_{K,L}$, for a system in the quantum state $\boldsymbol{K}$, to the state $\boldsymbol{L}$, due to any perturbation, is positive real, or zero, in value, and equal to $A_{L,K}$, the transition probability from $\boldsymbol{L}$ to $\boldsymbol{K}$. If M_K is the number of systems in $\boldsymbol{K}$, and $\dot{M}_K$ is the rate of change of this number with time, then

$$\dot{M}_K = -\sum_L A_{K,L} M_K + \sum_L A_{L,K} M_L = \sum_L A_{K,L}(M_L - M_K). \quad (8.3.1)$$

The rates of change, $\dot{M}_K$, must all be zero in the equilibrium ensemble. One solution is obviously that for which $M_K = M_L$ for all states, $\boldsymbol{K}$, $\boldsymbol{L}$. The probability W_K that a member chosen at random from the ensemble is in the state $\boldsymbol{K}$ is $W_K = M_K/M$, and is the same for all states $\boldsymbol{K}$. If Ω

is the total number of states between E and $E + \Delta E$, then

$$W_K = \Omega^{-1} \text{ (all } K \text{ for which } E \leqslant E_K \leqslant E + \Delta E). \qquad (8.3.2)$$

If all $A_{K,L}$ are non-zero, and hence positive, (2) is the only stationary solution for (1), since if the M_K's are not all equal there must be one $K = K^*$, for which

$$M_{K^*} \geqslant M_K \text{ (all } K\text{'s)}.$$

Equation (1) has then terms for $\dot{M}_{K^*}$ which are all zero or negative, and is negative unless $M_{K^*} = M_K$ for all K's.

It may be that the states K fall into two or more subsets, $K_\alpha, \ldots$, in subset α of Ω_α states, $K_\beta, \ldots$, in subset β of Ω_β states, such that for *all* pairs K_α, K_β,

$$A_{K\alpha,K\beta} = 0 \text{ (all } K_\alpha, K_\beta). \qquad (8.3.3)$$

In this case any value $W_{K\alpha} = x\Omega_\alpha^{-1}$ for all K_α, and $W_{K\beta} = (1 - x)\Omega_\beta^{-1}$ for all K_β, with $0 \leqslant x \leqslant 1$ will be stationary. The systems are then said to be non-ergodic, and to be ergodic if no separation into blocks $\alpha, \beta, \ldots$ of states can be made such that (3) holds. This is discussed more fully in the next section.

We have derived the following. The microcanonical ensemble for ergodic systems, corresponding to a stationary distribution of systems among the Ω allowed states, K, of the systems, is that in which there is an equal probability, $W_K = \Omega^{-1}$, for all quantum states K, that a member system of the ensemble, selected at random, will be in that state K.

From the fact that quantum states all correspond to the same volume, h^Γ, in the classical phase space $q^{(\Gamma)}$, $p^{(\Gamma)}$, it follows that the classical probability density, $W(p^{(\Gamma)}q^{(\Gamma)})$ must also be constant in the microcanonical ensemble. The direct demonstration from the classical equations of motion is as follows.

The equations of motion in the Hamiltonian form are

$$\dot{q}_i = \partial H/\partial p_i, \qquad \dot{p}_i = -\partial H/\partial q_i, \qquad (8.3.4)$$

for every pair, q_i, p_i, of the Γ conjugated coordinate and momenta components. Now consider all systems which have values $q^{(\Gamma)}$, $p^{(\Gamma)}$ of coordinates and momenta lying between q_i and $q_i + \Delta q_i$, p_i and $p_i + \Delta p_i$ for every i, that is those which are in the volume

$$h^\Gamma \Delta\Omega = \prod_{i=1}^{i=\Gamma} \Delta q_i \Delta p_i \qquad (8.3.5)$$

of phase space. At a later time these systems move to new coordinates and momenta determined by integrating (4). However, we shall show that the volume, $\Delta\Omega$, which just includes these systems, does not change with time, $\dot{\overline{\Delta\Omega}} = 0$. This statement is known as the Liouville theorem.

From (5) we have

$$\dot{\overline{\Delta\Omega}} = \Delta\Omega \sum_i (\Delta q_i^{-1}\, \dot{\overline{\Delta q_i}} + \Delta p_i^{-1}\, \dot{\overline{\Delta p_i}}), \tag{8.3.6}$$

from (4),

$$\begin{aligned} \dot{\overline{\Delta q_i}} &= \dot{\overline{q_i + \Delta q_i}} - \dot{q}_i \\ &= \partial H(q_i + \Delta q_i)/\partial p_i - \partial H(q_i)/\partial p_i \\ &= \Delta q_i(\partial^2 H/\partial p_i \partial q_i), \end{aligned} \tag{8.3.7}$$

and

$$\dot{\overline{\Delta p_i}} = -\Delta p_i(\partial^2 H/\partial p_i \partial q_i), \tag{8.3.8}$$

so that each member of the sum in (6) is zero,

$$\Delta q_i^{-1}\, \dot{\overline{\Delta q_i}} + \Delta p_i^{-1}\, \dot{\overline{\Delta p_i}} = 0, \tag{8.3.9}$$

and hence the sum itself is zero.

Now it follows that if $W(\boldsymbol{p}^{(\Gamma)}, \boldsymbol{q}^{(\Gamma)}, t)$ is given at time $t = 0$, its numerical value at a certain point $\boldsymbol{p}_0^{(\Gamma)}, \boldsymbol{q}_0^{(\Gamma)}$, of phase space at $t = 0$ will later be found at t at the point $\boldsymbol{p}^{(\Gamma)}(\boldsymbol{p}_0^{(\Gamma)}, \boldsymbol{q}_0^{(\Gamma)}, t)$, $\boldsymbol{q}^{(\Gamma)}(\boldsymbol{p}_0^{(\Gamma)}, \boldsymbol{q}_0^{(\Gamma)}, t)$ to which a system originally at $\boldsymbol{p}_0^{(\Gamma)}, \boldsymbol{q}_0^{(\Gamma)}$ will have arrived at time t. If W at a fixed point in phase space is not to change with time, the density W must be constant along the path in phase space which the systems traverse. The system is said to be ergodic if it traverses all phase space before returning to its original position, or more carefully stated, if it comes within a distance ε of all points, where ε is infinitesimal before returning to within a distance ε of its original position. In this case only a uniform density is static in time.

8.4. The ergodic hypothesis

The ensemble representing the totality of all experimental systems prepared with fixed V, E, N has equal probability of all states, $\boldsymbol{K}$, if the systems are ergodic.

A truly isolated single system is non-ergodic, indeed this is exactly the meaning of the term "stationary state," namely that the quantum state K labels a function Ψ_K which is the solution of the time independent of Schrödinger equation, $\mathscr{H}\Psi_K = E_K\Psi_K$, and properties computed from this function are independent of time. A system once in this state will not undergo transitions to other states, except under the influence of some perturbation.

However, the systems which one constructs in the laboratory with supposedly fixed V, E, N do not correspond to this idealization. Firstly, there are at least microscopic differences in the container walls for the different systems. Even the single system is actually subject to small time dependent fluctuations in the Hamiltonian due to the surroundings. The states K that one computes for the members of the ensemble are those with the fictitious Hamiltonian of an idealized completely isolated system. This Hamiltonian is to be chosen with a potential energy term appropriate to a statistical time average of the walls of the real systems. The states of the real systems can then always be represented as linear combinations of the stationary states of this idealization, but the coefficients will be time dependent. The differences between the container walls of the real systems and that of the idealization, and the time dependent fluctuations in the real walls, are the perturbations that cause the transition probabilities A_{KL} of (8.3.1) to differ from zero.

The question remains as to what transition probabilities are actually induced in the real systems.

Actually in one sense no real system is ever ergodic. For any system in a chemical laboratory, at the energy corresponding to some temperature encountered in laboratory procedure, there exist conceivable quantum states, which would be reached by the same contents of the system were it first "heated" by the addition of much more energy and then cooled to the original energy, but to which no transitions occur at the energy specified. For instance quantum states corresponding to differing numbers of atoms of the atomic elements, which could be reached by nuclear transitions, are always excluded. In general the specification of the set $N = N_a, N_b$, of chemical components of the system is not trivial. In the ensemble corresponding to an ordinary laboratory system this set at least specifies the number of each isotope of each element present, rather than the total number of nucleons, which latter specification might be correct for the same matter under the conditions present in the interior of some stars. The laboratory system is certainly non-ergodic with respect to nuclear reactions.

In very many, if not most, of the laboratory systems of interest, it is necessary to be more specific than this, and to specify the number of molecules of each molecular species. Certain chemical reactions must be assumed to proceed at zero rate. For instance, a mixture of H_2, O_2 and H_2O at room temperature in the absence of a catalyst is essentially stable. It is a perfectly suitable system for the application of thermodynamic laws, although true equilibrium with respect to the chemical reaction $H_2 + \frac{1}{2}O_2 \rightarrow H_2O$ is not present. The corresponding ensemble would be one constrained to those quantum states having fixed values of N_{H_2}, N_{O_2} and N_{H_2O}. At higher temperatures, or in the presence of a catalyst, equilibrium with respect to this reaction would be attained, and the restriction should be made only to states of fixed numbers of the two kinds of atoms present, M_H and M_O, including all stoichiometric allowable numbers of molecules. Practically no system containing specified numbers of organic molecular species is at equilibrium with respect to all reactions.

Now the type of requirement in the specification of the system is completely familiar in thermodynamics. The specification of a thermodynamic state is not always trivial. One must specify which of the conceivable chemical reactions are assumed to have proceeded to equilibrium, and which are assumed to be inhibited to zero rates. Exactly the same requirement is placed on the specification of the allowed quantum states of the ensemble.

One must not assume, however, that only the occurrence or nonoccurrence of chemical reactions need by specified. In many cases two or more entirely different thermodynamic phases of the same species may be able to exist (at least metastably) at one temperature and pressure, and even to coexist in intimate contact. The specification: Avogadro's number of carbon atoms at $P = 1$ atmosphere, $T = 300°K$, requires the additional information that the phase is pure graphite, or if not, what fraction is present as diamond impurity. Similarly a sample of H_2 gas at 30°K in a non-catalytic container requires the specification of the fraction of ortho-hydrogen, namely of the relative fraction in even or odd rotational states.

In general the experimentalist in preparing a system in a given thermodynamic state, must investigate whether all systems prepared in this state have the same properties. If they differ he seeks a new thermodynamic variable, x, whose specification now assures reproducibility. In the examples discussed this x may be: the number N_{H_2O} of water molecules if previously only M_H and M_O were specified, the fraction of carbon present in the diamond phase, or the fraction of H_2 molecules present as ortho-hydrogen. The Ω quantum mechanical states of the ensemble corresponding to the original state prescription will fall into differing sets, one set consisting of Ω_x states corresponding to the new thermodynamic variable x having the value x, and $\Omega = \sum_x \Omega_x$. The fact that the new variable x can be experimentally fixed and held constant means that the experimental ensemble is non-ergodic unless x is specified, and the correct ensemble for fixed x has Ω_x allowed states.

Obviously no completely general mathematical considerations, without detailed knowledge of the system concerned, can inform us whether a particular idealized mathematical model, assumed to be ergodic within all states corresponding to a fixed set, $x_1, x_2, \ldots, x_n$, of n macroscopic variables will be a true representation of the experimental systems or not. In certain cases, however, and that of ortho and para-hydrogen is an example, the detailed examination of the mechanical states can make clear the interpretation of what might otherwise be a completely mysterious variable x.

The question still remains whether there are non-ergodic systems of another class. Assume that there exist sets, Ω_α, ${}_\beta\Omega$, . . ., of states between which no transitions occur at the temperature, density, and composition for which we calculate. Assume that the properties of the overwhelming majority of the states ever occupied by systems of the laboratory preparation, even if belonging to differing values of α, have indistinguishable properties. We will then not detect a new thermodynamic variable x corresponding to α. The answer is that such cases are not uncommon. In such an example we must use for Ω the sum $\Omega = \Sigma\Omega_\alpha$ for all α's actually reached in the laboratory preparation. The known examples of this type are all cases in which a system is "supercooled" or "frozen" into a thermodynamically metastable state. It is prepared from a state at higher energy in which the system is ergodic with respect to transitions between states differing in the physical characteristic corresponding to the subscript α. Examples include all glasses, many solid solutions, and a few one-component crystals such as ice and CO. We shall discuss the example of carbon monoxide.

The molecule CO, which is isoelectronic with N_2, is very nearly symmetrical with respect to end-to-end exchange. At the melting point the various configurations of the crystal which differ in mutual orientations "carbon up" or "oxygen up", have so little difference in energy that equilibrium consists of all 2^N possible mutual orientations in practically equal amounts. As the crystal is cooled towards 0°K the actual configuration in any one crystal becomes frozen, the transitions to other configurations become forbidden and the system is non-ergodic. This occurs above the temperature at which the presumable single triply periodic mutual orientation of minimum energy is appreciably weighted at equilibrium. For the overwhelming majority of the random configurations the properties are indistinguishable. In this case, the ensemble corresponding to the laboratory preparation consists of 2^N differing configurations. Each of these 2^N configurations includes essentially the same states of vibration in the crystal.

Anticipating our later result that the entropy, S, is given by $S = k \ln \Omega$, the value approached as $T \rightarrow 0°\text{K}$ is $S = kN \ln 2$, instead of $S = 0$. This is the observed experimental behavior.

In this case, if another method of preparation could be found, either by some exotic experimental technique, or by sufficiently long annealing at the appropriate temperature, such that the equilibrium mutual orientation were actually obtained, the thermodynamic properties would be different. The ensemble would then be that in which one (or a vanishingly small fraction) of the 2^N mutual orientations, was alone present. This single configuration would have a significantly lower energy. At fixed total energy more energy would be available for the vibrational states of the crystal, and hence more vibrational states available to this one configuration than to all the others combined. If the annealing were such that only partial equilibrium were attained the apparent thermodynamic properties would depend on the length of the annealing process. We would have no easy method of assigning the correct ensemble, but correspondingly the system would not be considered suitably defined in the thermodynamic sense. Many such cases exist.

This example emphasizes that the entropy is not only a property of the particular system, but also of our knowledge of the system. The single non-ergodic crystal at low temperatures has a single, although disorderly, array of orientations. It is, however, only a single specific member of a class, the ensemble, that includes $2^N - 1$ other members which we are unable to distinguish from it. This indistinguishability increases its entropy, both in our formalism, *and* in values obtained by classical calorimetric methods.

This example may be understood more intuitively by considering the hypothetical equilibrium between crystal and vapor. At equilibrium the rate of sublimation of surface molecules into the vapor, and the rate of condensation of vapor molecules are equal. Each molecule on the surface, having a fixed orientation, has a certain probability of vaporization per unit time, and the flow into the vapor is the number of surface molecules times this probability per unit time. The return flow is equal. Now if a Maxwell Demon has a prearranged table of the orientations at each surface site, no matter how disorderly, and returns to the vapor the half of the molecules that attempt to condense in the orientation contrary to his table, the return flow will be halved at the same pressure. The equilibrium vapor pressure must be doubled to reattain equality of the two rates. But a change $\Delta S = -R \ln 2$ per mol in the entropy of the crystal requires just double the vapor pressure at equilibrium. Thus the crystal of single orientation, if this orientation is completely defined and prescribed by the Maxwell Demon, has the lower entropy by $R \ln 2$ per mole than has the assembly of all disordered arrangements.

The example is, of course, unreal, not only because Maxwell Demons are rare, but also because an appreciable rate of sublimation and recondensation would be a mechanism for making transitions between crystals of different orientations. The ensemble would no longer be non-ergodic. If one completely regular orientation were more stable at the energy described, then this one, with lower configurational entropy, but higher entropy in the lattice vibrations (and hence corresponding to a higher temperature if the energy were fixed) would be formed.

To summarize, then, the microcanonical ensemble is one of equal probabilities of those quantum states reached by the experimental systems we wish to discuss. Sometimes the systems may be non-ergodic under certain conditions, such as that of energies corresponding to low temperatures, for which we wish to compute. If, however, the experimental systems are prepared under conditions for which they are ergodic we include all quantum states reached in their preparation.

8.5. The quantity $S = k \ln \Omega$

The number of quantum states, Ω, available to a system of fixed V, E, N, and with fixed values of any other thermodynamic variables necessary to describe the thermodynamic state uniquely will be proportional to ΔE if the energy is prescribed to lie between E and $E + \Delta E$. We use $\Omega(V,E,N)$ for the number of states per unit range of energy, so that

$$\Omega = \Omega(V,E,N)\Delta E, \tag{8.5.1}$$

and define

$$S = k \ln \Omega. \tag{8.5.2}$$

From (1), we have

$$S = k \ln \Omega(V,E,N) + k \ln \Delta E. \tag{8.5.3}$$

Now actually, for purposes of computation, the last term, $k \ln \Delta E$ is always dropped as numerically negligible, which is equivalent to choosing $\Delta E = 1$ in eq. (1). This choice, of course, means that Ω and S depend on the energy units used.

That this neglect is really justifiable is evident after we have identified S with entropy, and remember that the entropy is of order kN, so that the fractional error is of order $N^{-1} \ln \Delta E$. For macroscopic systems with N of order 10^{20} or greater, the fractional error is less than 10^{-17} unless ΔE exceeds $\exp 10^3 \simeq 10^{430}$, or is less than 10^{-430}. Even if the quantum mechanical uncertainty limit for a system isolated over the lifetime of the universe, $t = 10^{10}$ years $= \pi \times 10^{17}$ sec, $\Delta E \geqslant h/t$, is used, one has $\Delta E \geqslant 10^{-44}$ ergs, far closer to unity than our limit. The largest uncertainty that we may permit must be sufficiently small to identify the thermodynamic state reasonably, and hence less than the energy of the system, of order $NkT \simeq 10^7$ ergs for room temperature and $N = 10^{20}$. We are scarcely likely to use units that bring ΔE into a dangerous range; even megavolts $= 10^{18}$ ergs/mol would be completely permissible.

We wish to emphasize that eq. (3) which writes S as dependent on the experimental uncertainty, ΔE, in the energy of the system is logically necessary. As we discussed before in section 8.4 the entropy depends on the knowledge that we have of the system. Greater precision of knowledge always leads to a lower value of S. For numerical purposes the additive $k \ln \Delta E$ may be neglected, but it logically belongs in the definition of S.

Some authors prefer to define Ω as

$$\Omega = \int_{-\infty}^{E} -\Omega(V,E',N)dE'$$

so that Ω is the total number of states of energy below E, i.e. $E_K \leqslant E$. For most systems this differs numerically negligibly from (1). However, we find this to be illogical. More practically it requires that the quantity $T^{-1} = (\partial S/\partial E)_{N,V}$, which we show (section 8.6) to be the reciprocal temperature is *always* positive. Normally this is the case with the definition (1), since, if all possible degrees of freedom are included $\partial\Omega(V,E,N)/\partial E > 0$. However, at least one case of spin energy has $\Omega(E,N)$ increasing to a maximum at a certain energy, and then decreasing as E increases further, so that the reciprocal temperature can take negative values for this peculiar degree of freedom. The negative reciprocal temperature range has been reached experimentally.†

If two separately isolated systems, a and b, of value $S_a = k \ln \Omega_a$ and $S_b = k \ln \Omega_b$, are considered as a single system (ab), the value of the quantity $S_{(ab)}$ defined by (2) for the combination is

$$S_{(ab)} = k \ln \Omega_{(ab)} = S_a + S_b. \tag{8.5.4}$$

This follows from the fact that the quantum state of the combined system (ab) is given by specifying the state in both subsystems, a and b,

$$\boldsymbol{K}_{(ab)} = \boldsymbol{K}_a, \boldsymbol{K}_b. \tag{8.5.5}$$

The number of states of the combined system is then the product of the number in a and b,

$$\Omega_{(ab)} = \Omega_a \Omega_b, \tag{8.5.6}$$

since every state $\boldsymbol{K}_a$ of a is permissible with each of the Ω_b states of b. From this, (4) follows.

The quantity S also has the property of increasing, (or remaining effectively unchanged) for every process permitted to occur in an isolated system of constant total energy E. This follows from the interpretation in terms of quantum states of what is meant by permitting a thermodynamic change to occur in a system. Such a change always corresponds to permitting the system to occupy new quantum states, corresponding to new values of some macroscopic variable, hitherto forbidden to the system.

For instance, consider the change represented by the occurrence of a chemical reaction such as $H_2 + \frac{1}{2}O_2 \rightleftarrows H_2O$. The initial state of the system is specified by the numbers, $N_{H_2}{}^{(o)}$, $N_{O_2}{}^{(o)}$, $N_{H_2}{}^{(o)}$, of the three molecular species. The system is non-ergodic with respect to quantum states having different numbers of these molecules. Permitting the reaction to proceed, say by insertion of a catalyst, permits new quantum states to occur: those characterized by any set,

$$\begin{aligned} N_{H_2} &= N_{H_2}{}^{(o)} - 2N, \\ N_{O_2} &= N_{O_2}{}^{(o)} - N, \\ N_{H_2O} &= N_{H_2}{}^{(o)} + 2N, \end{aligned}$$

---

† E. M. Purcell and R. V. Pound, *Phys. Rev.* **81**, 279 (1951).

as long as all numbers remain positive or zero. Since the number, Ω, of allowed states increases, its logarithm, and S, also increase.

Other examples include a combined system (ab) of two parts a and b of fixed V_a, E_a, N_a, V_b, E_b, N_b, with total $V = V_a + V_b$, $E = E_a + E_b$ and $N = N_a + N_b$. Any change at fixed V, E, N, such as that of thermal contact for which $\Delta E_a = -\Delta E_b$, or the moving of a wall separating the two parts for which $\Delta V_a = -\Delta V_b$, or diffusion between them for which $\Delta N_a = -\Delta N_b$, allows new quantum states, and Ω and S increase.

We now show that *at equilibrium* the restriction of a system to forbid such changes results in only a negligible and non-measurable decrease of S. For instance: removing the catalyst of a chemical reaction; removing the thermal contact between two parts, a and b; the stopping of motion of a wall separating a and b; or closing a stopcock which prevents further diffusion. If the inhibition is imposed at equilibrium, it results in no measurable decrease of S. In all such cases there is some variable, α, which might be the extent of the reaction, N, the changes ΔE_a, ΔV_a, or ΔN_a of the energy, volume, or number set in part a. This α measures the distribution in the system, and specifies the particular set of Ω_α quantum states permitted the system.

The total number, Ω, of states permitted the uninhibited system is

$$\Omega = \sum_{\alpha=1}^{\alpha=M} \Omega_\alpha, \tag{8.5.7}$$

where M is the total number of distributions, that is the total number of values that α can take. The probability, $W(\alpha)$, of finding a member of the uninhibited ensemble, after equilibrium has been reached, in the distribution α is

$$W(\alpha) = \Omega_\alpha/\Omega. \tag{8.5.8}$$

There is one value of α, namely $\alpha = m$ for which $W(m)$ is a maximum

$$W(m) \geqslant W(\alpha), \qquad \Omega_m \geqslant \Omega_\alpha \text{ (all } \alpha). \tag{8.5.9}$$

We may bracket $S = k \ln \Omega$ in value between

$$k \ln \Omega_m \leqslant S = k \ln \Omega \leqslant k \ln \Omega_m + k \ln M$$

$$S_m \leqslant S \leqslant S_m \left(1 + \frac{\ln M}{\ln \Omega_m}\right). \tag{8.5.10}$$

We again make use of the fact that $\ln \Omega_m$ is of order N, the number of molecules of the system, and M, the total number of distributions is also of this order. The correction term on the right of eq. (10) is then of order $N^{-1}\ln N$, which, for a macroscopic system of $N \geqslant 10^{20}$ is indeed numerically negligible. It follows that the value of the quantity S for the uninhibited system is numerically equal, with negligible error, to that for the inhibited system in the most probable distribution.

Several comments appear to be required. Firstly, we have described α as an integer, which it will be for a chemical reaction, or for the case of diffusion. In the case of a continuous variable ΔE_a or ΔV_a of the example given we must assign an experimental uncertainty δE_a or δV_a and use α as the number of units of this uncertainty, $\Delta E_a = \alpha\delta E_a$, or $\Delta V_a = \alpha\delta V_a$, to have a denumerable M for the total number of distributions in eq. (10). In any case if no uncertainty principle is violated in the choice of δE_a or δV_a the quantity $N^{-1}\ln M$ will indeed be negligible for $N \geqslant 10^{20}$.

Secondly, we may, on reimposing the inhibition after equilibrium has been attained with probability $W(\alpha) < W(m)$ trap a single system of the ensemble in a state α different from the most probable one. The value of $S_\alpha = k\ln\Omega_a$ will then be less than $S \simeq S_m = k\ln\Omega_m$ by the amount,

$$\Delta S_\alpha = S_m - S_\alpha = k\ln(\Omega_m/\Omega) = k\ln[W(m)/W(\alpha)],$$

$$W(\alpha) = W(m)\exp-(\Delta S_\alpha/k). \tag{8.5.11}$$

However, if ΔS_α is numerically measurable, say $\Delta S_a = 10^{-6}S_m \simeq 10^{-6}kN$ for $N = 10^{20}$, the probability of this happening is of the order exp (-10^{14}). This would indeed be an unlikely event.

Thirdly, if an experimenter with the characteristics of a Maxwell Demon is able to inhibit the distribution of a system along some 10^{20} different components, $\alpha_1, \alpha_2, \ldots, \alpha_\nu, \ldots, a_{10}{}^{20}$, with α_ν taking M_ν values, then S will be decreased by the amount

$$S = -\sum_{\nu=1}^{\nu=10^{20}} k\ln M_\nu$$

by the inhibition, even at equilibrium. Such an inhibition to a known specified distribution along a number of degrees of freedom approaching in magnitude the number of microscopic degrees of freedom will indeed decrease S. It is essential to our discussion that we deal with ensembles of systems for which the number of *macroscopic* degrees of freedom, the thermodynamic distribution variables, is negligible compared with the number of *microscopic* mechanical degrees of freedom.

Finally we may note that, in the case of diffusion of a single species of molecule between two parts a and b, the quantum states with fixed values of N_a indistinguishable identical molecules in part a, and N_b in part b, are not different for differing distributions of any artificial numbering of the identical molecules. Were the states different depending on which *numbered* molecules were in a, so that a distribution were determined by giving the location, in part a or part b, for each of the $N = N_a + N_b$, numbered molecules, the number, M, of distributions would have been $M = 2^N$, and $\ln M = N \ln 2$ would not be negligible compared with N. Since the classical phase space requires identification of the coordinates, it is different for positions with molecule i in a and j in b, and those in which these positions are exchanged. Before the recognition that a quantum mechanical state function must be symmetric or antisymmetric in these permutations, and hence not described by an artificial numbering of identical particles, it was recognized by Gibbs to be necessary to divide the classical phase space by N! This was necessary in order to make S an extensive property.

Since S is not changed by imposing an inhibition at equilibrium it is unchanged by separating a system into two unconnected parts, a and b, at equilibrium. We have already seen that $S_{(ab)}$ for a combined system (ab) of two isolated parts, a, and b, is the sum of S_a and S_b, eq. (4). It follows that the quantity S for a system of uniform intensive properties, P, T, composition, and phase, etc., is proportional to the size of the system.

The quantity S is extensive. It increases, or does not measurably change, for all thermodynamic changes at constant V, E, N.

8.6. The quantities T and P

We define T and P by the equations

$$T^{-1} = (\partial S/\partial E)_{V,N} \tag{8.6.1}$$

$$P = T(\partial S/\partial V)_{E,N}. \tag{8.6.2}$$

Consider two-part systems, a and b, each of fixed volume, V_a and V_b, of fixed energies E_a and E_b, respectively, and each with fixed content, N_a, N_b. Permit thermal contact between a and b, so that energy may flow from b to a, but with $E = E_a + E_b$ held constant, so that $dE_a = -dE_b$. The equilibrium values of E_a, E_b, will be those for which the number

$$\Omega(V_a,E_a,N_a;\ V_b,E_b,N_b) = \Omega_a(V_a,E_a,N_a)\Omega_b(V_b,E_b,N_b) \tag{8.6.3}$$

of quantum states is a maximum, or $S = k \ln \Omega = S_a + S_b$ is maximum.

Writing

$$\delta S = \delta S_a + \delta S_b = (\partial S_a/\partial E_a)\delta E_a + (\partial S_b/\partial E_b)\delta E_b \geqslant 0, \qquad (8.6.4)$$

and using $\delta E_a = -\delta E_b$, with (1) defining T,

$$\delta S = (T_a^{-1} - T_b^{-1})\delta E_a \geqslant 0. \qquad (8.6.5)$$

The change will go in such a direction that energy flows from b to a, $\delta E_a > 0$, if $T_a < T_b$, and in the reverse direction if $T_b > T_a$. Heat flows from high T values to low T values, whatever the nature of the systems, and equilibrium is established if $T_a = T_b$. It follows that T is some scale of temperature.

The quantity P of eq. (2) is seen, from (1) to have the dimension of E/V, which is the dimension of a pressure.

Again consider two-part systems, a and b, at equilibrium with respect to heat flow, so that $T_a = T_b = T$, and with fixed $V = V_a + V_b$, but let a wall separating them move so that $dV_a = -dV_b$. Again write, using (1) and (2), that

$$\begin{aligned}\delta S = \delta S_a + \delta S_b &= (T_a^{-1} - T_b^{-1})\delta E_a + T^{-1}(P_a - P_b)\delta V_a \\ &= T^{-1}(P_a - P_b)\delta V_a \geqslant 0. \qquad (8.6.6)\end{aligned}$$

The change will go so that the system of higher P value gains in volume at the expense of that of the other, and equilibrium is established when $P_a = P_b$. Two systems that are at equilibrium with respect to such a volume exchange have equal mechanical pressures. We have established that equality of the values of P for two systems in thermal equilibrium means that they have equal pressures.

One easy procedure from here would be to compute P from eq. (2) for a perfect gas and show that it is equal to the mechanical pressure. Since for one system P is actually the pressure it follows from what we have proved that it is the pressure for all systems.

Instead we use a more abstract derivation that proceeds directly from (2) and (1) without using the demonstration (6). For changes at constant N (and any other extensive variables necessary to define the system) we have from (1) and (2),

$$(dS)_N = T^{-1}(dE)_N + (P/T)(dV)_N, \qquad (8.6.7)$$

or

$$P = -(\partial E/\partial V)_{S,N}, \qquad (8.6.8)$$

$$(dE)_{S,N} = -PdV. \qquad (8.6.9)$$

The change in energy of a system of fixed S and N is $-PdV$.

Now consider the ensemble of systems of fixed V, N consisting of an equal number of systems in each of the $\Omega = \Omega(V,E,N)\Delta E$ quantum states of energy between E and $E + \Delta E$. The average change of energy of the members of this ensemble upon changing V by dV is

$$\langle \partial E/\partial V\rangle_{\text{Av}} = \Omega^{-1} \sum_{K=1}^{K=\Omega} (\partial E_K/\partial V), \tag{8.6.10}$$

namely the average of the change in each quantum state K. The average mechanical pressure, P_{mech}, is just $P_{\text{mech}} = -\langle \partial E/\partial V\rangle_{\text{Av}}$, but $\langle \partial E/\partial V\rangle_{\text{Av}}$ of (10) is just the change in energy of the systems in the fixed Ω quantum states, namely $(\partial E/\partial V)_{S,N}$ of (8). Hence P defined by (2) is actually the mechanical pressure.

8.7. The thermodynamic laws

It is essential that we make the initial assumption that on the microscopic scale in which we identify the mechanical coordinates and momenta of all the component parts of the systems there are only conservative forces. In other words no frictional forces exist at the microscopic level. For an isolated system with conservative forces the Hamiltonian is constant in time, energy is conserved. This is the first law of thermodynamics.

In the previous two sections a quantity S related to the pressure P, and a quantity T which has the qualitative characteristics of temperature, has been defined. The relations are those which define the thermodynamic entropy. The quantity S is extensive, and increases for all processes that may occur in an isolated system of constant V, E, N. This is the second law, the differential relations between S, P, T being such that

$$dE = TdS - PdV,$$

where T is the absolute thermodynamic temperature, therefore TdS is the heat flow for reversible processes when PdV is the only work done.

The entropy S is k times the logarithm of an integer, the lowest value of which is unity when the systems are constrained by the macroscopic conditions to a single quantum state. For many real systems, and indeed for all which are truly ergodic at low enough temperatures, the degeneracy of the quantum level of lowest energy is a small number independent of the size of the system, and S approaches zero in numerical value as E approaches its minimum. For these systems the zero of entropy is actually approached as T goes to zero. There remain a number of systems which become non-ergodic on cooling (such as CO, section 8.4) and do not approach zero entropy at $T = 0°\text{K}$.

These, then establish the thermodynamic laws or equilibrium systems, with the third law left in the somewhat ambiguous position of being only correct when it is correct, namely for systems which remain ergodic down to absolute zero.

However, we have avoided discussing one property of real systems. This would require the more complicated discussion of systems in which the macroscopic state changes with time. We have shown that if a change is permitted in an isolated system the average member of the ensemble will, after infinite time, have an increase in S. We have *not* shown that this increase proceeds monotonically, and that an examination after a finite time interval will never show a decrease in S. This demonstration would be necessary to completely derive the laws of thermodynamics. The demonstration is not undertaken in this book, but is discussed briefly in section 8.10.

8.8. The canonical ensembles

We now derive the probability expression for quantum states of a system whose thermodynamic state is defined by alternative variables other than V, E, $\mathbf{N}$. The variables V, E, $\mathbf{N}$ are all extensive, and to each there is a conjugate intensive variable, $-P$ for V, T for E, and μ_a for any component number, N_a, of $\mathbf{N}$. The thermodynamic state may equally well be prescribed by replacing any extensive variable with its conjugate, providing only that at least one extensive variable remains to define the size of the system.

We consider an ensemble of supersystems, but use this ensemble only in order to talk about the average properties of parts of the supersystem. The supersystem has fixed total V_t, E_t, and $\mathbf{N}_t$, but consists of $M + 1$ connected parts. One unique part we call the reservoir, the other parts are the M systems in which we are interested. We wish, now, to discuss the properties of the ensemble of these systems which are in contact with the reservoir, and through the reservoir with each other.

In particular consider a distribution for one supersystem characterized by a number, M_K, giving the number of the systems in a particular quantum state $\mathbf{K}$ of fixed V, $\mathbf{N}$, and energy E_K. For fixed value of M_K the total number of quantum states, $\Omega_t(M_K)$ of the supersystem will depend on M_K. It will be the product of the number of ways, $M!/(M - M_K)!M_K!$ that we can pick the particular M_K systems out of the M different systems, times a number, $\Omega_r(M_K)$. This Ω_r is the number of states allowed to the reservoir plus the $M - M_K$ systems in states other than K. This reservoir plus $M - M_K$ systems has volume $V_t - M_K V$, energy $E_t - M_K E_K$, and a number set $N_t - M_K N$. The quantity $S_r = k \ln \Omega_r$ is the entropy of such a complex. Writing $S_t(M_K) = k \ln \Omega_t$ we have

$$S_t(M_K) = k \ln \frac{M!}{(M - M_K)!M_K!} + S_r(V_t - M_K V, E_t - M_K E_K, N_t - M_K N). \quad (8.8.1)$$

The equilibrium value of $M_K = \bar{M}_K$ is such that $S(\bar{M}_K)$ is a maximum

$$dS_t/d\bar{M}_K = 0 = k \ln \frac{M - \bar{M}_K}{\bar{M}_K} + (dS_r/d\bar{M}_K). \quad (8.8.2)$$

We can truly neglect $\bar{M}_K$ compared with M in the numerator of the logarithm term. Write $W_K = \bar{M}_K/M$ at equilibrium, for the probability that a system of the ensemble will be in state K. Use

$$(\partial S_r/\partial E)_{V,N} = T^{-1}, \ (\partial S_r/\partial V)_{E,N} = P/T, \text{ and } (\partial S_r/\partial N_\alpha)_{V,E,N_\beta} = -\mu_\alpha/T$$

to find

$$W_K = \exp \frac{1}{k} \frac{dS_r}{dM_K} = \exp \frac{1}{kT}[-PV + N \cdot \boldsymbol{\mu} - E_K]. \quad (8.8.3)$$

The use of the differential expressions such as $(\partial S/\partial E)_{V,N} E_K$ for the change in S_r due to the energy removed per unit increase of M_K is valid only if E_K is negligible compared with E_t. Similarly V and N must be negligible compared with V_t and N_t, respectively. Our equation (3) is asymptotically correct only in the limit that the supersystem approaches infinite size.

Now so far we have been intentionally vague about what characterized the ensemble of M systems in each supersystem. It is this difference of characterization that leads us to different ensembles.

The first case we may choose is a trivial one in which there is, actually, no contact between the systems and the reservoir, and the M systems are all those having fixed volume V, containing the set N of molecules, with energies between E and $E + \Delta E$. With $E_K = E$ the quantity in brackets in the exponent on the right of (3) is

$$-PV + N \cdot \boldsymbol{\mu} - E = -PV + G - E = -TS, \quad (8.8.4)$$

and the probability, W_K, is the same for all states, K,

$$W_K = e^{-S/k}. \tag{8.8.5}$$

Since W_K is the same for all states it must be the reciprocal of the number, Ω, allowed, and we recover our old equation, $S = k \ln \Omega$.

The second case we may choose is that in which the walls of the systems are rigid and impermeable to molecules, all systems have fixed V and N, but the walls conduct heat and all energies E_K are permitted. Since $-PV + G = A$ we have

$$W_K = \exp{(A - E_K)}/kT \tag{8.8.6}$$

for states K of fixed V, N, but any energy E_K. The sum of W_K over all allowed K values must be unity, so that

$$A = -kT \ln Q_c(V,T,N), \tag{8.8.7}$$

$$Q_c(V,T,N) = \sum_K e^{-E_K/kT}. \tag{8.8.8}$$

The quantity Q_c is the partition function for the petit canonical ensemble of fixed V, T, N. Since $\Omega(V,E,N)$ is defined as the number of states per unit energy range at E one may write (8) as

$$Q_c(V,T,N) = \int_0^\infty \Omega(V,E,N)\, e^{-E/kT} dE. \tag{8.8.9}$$

The third case is that the systems are confined to the volume V by rigid walls, but these are permeable to all the molecular species as well as to energy. States K for all values of N as well as all E_K occur. Setting the sum of W_K equal to unity we find

$$-PV = -kT \ln Q_{G.C}(V,T,\mu), \tag{8.8.10}$$

$$Q_{G.C}(V,T,\mu) =$$

$$\sum_N e^{\mu \cdot N/kT} Q_c(V,T,N), - \sum_N e^{\mu \cdot N/kT} \int_0^\infty \Omega(V,E,N)\, e^{-E/kT} dE. \tag{8.8.11}$$

The partition function $Q_{G \cdot C}$ is that of the grand canonical ensemble.

Obviously we could consider cases between the second and third in which the walls were permeable to heat and to molecular species α, β, γ, . . . but not to species a, b, ... The number set $N^{(a)} = N_a$, N_b, . . . is then fixed in all systems, but the set $N^{(\alpha)} = N_\alpha$, N_β, . . ., must be summed over all values. Again setting the sum of W_K equal to unity we have

$$[\mathbf{N}^{(a)}\cdot\boldsymbol{\mu}^{(a)} - PV] = -kT \ln \sum_{\mathbf{N}^{(\alpha)}} \exp \frac{\mathbf{N}^{(\alpha)}\boldsymbol{\mu}^{\cdot(a)}}{kT} Q_c(V,T,\mathbf{N}^{(a)},\mathbf{N}^{(\alpha)}), \tag{8.8.12}$$

which goes over to (7) if there are no species of type α, and to (10) if there are none of type a. The thermodynamic function

$$\Theta_a = \mathbf{N}^{(a)},\, \boldsymbol{\mu}^{(a)} - PV = \Theta_a(V,T,\mathbf{N}^{(a)},\, \boldsymbol{\mu}^{(\alpha)}) \tag{8.8.13}$$

is a "natural" function of $V,T,\mathbf{N}^{(a)}$, $\boldsymbol{\mu}^{(a)}$, namely,

$$\Theta_a = \mathbf{N}\cdot\boldsymbol{\mu} - PV - \mathbf{N}^{(\alpha)}\cdot\boldsymbol{\mu}^{(\alpha)} = A - \mathbf{N}^{(\alpha)}\cdot\boldsymbol{\mu}^{(\alpha)}, \tag{8.8.14}$$

$$d\Theta_a = -PdV - SdT + \sum_a \mu_a dN_a - \sum_\alpha N_\alpha d\mu_\alpha. \tag{8.8.15}$$

In the limit that no species a exist $\Theta_0 = -PV$ which is a natural function of $V,T,\boldsymbol{\mu}$.

Another interesting case is that in which we permit the walls of the systems to move, so that the volumes of the systems may fluctuate. We keep the walls impermeable to all molecular species, but the walls conduct heat. The states $\mathbf{K}$ that occur are those of fixed N, but with all volumes and energies, and again we set the sum of all W^K equal to unity.

However, here a caution is necessary. The quantity $\Omega(V,E,\mathbf{N})$ is the number of states per unit of energy range of a system of fixed set $\mathbf{N}$ of molecules confined by rigid walls of volume V. It sums all states in which the molecules are confined within V, *including those in which the molecules actually occupy only some part of the volume* V. This is most obvious if we consider the classical limit that Ω is given by an integral

$$\Omega = \lim_{\Delta E\to 0} \frac{1}{\Delta E}\frac{1}{N!} \iint\limits_{V} \cdots \int_{E\leqslant H(p^{(N)}q^{(N)})\leqslant E+\Delta E} (d\mathbf{q}^{(N)}d\mathbf{p}^{(N)}/h^{\Gamma}) \tag{8.8.16}$$

which obviously includes the configurations in which all molecules huddle in one part of the volume V.

If the walls of the systems are movable, acting under the pressure, P, of the reservoir, they will collapse to the actual volume occupied by the molecules. The number of states for volumes between V and $V + \Delta V$

and between E and $E + \Delta E$ will then be $[\partial\Omega(V,E,N)/\partial V]\Delta V \Delta E$. The requirement that $\Sigma W_K = 1$ then gives

$$G = -kT \ln \int\int_0^\infty [\partial\Omega(V,E,N)/\partial V]\left\{\exp -\frac{1}{kT}(PV+E)\right\} dV dE. \tag{8.8.17}$$

By a partial integration, since Ω for $V \to 0$ approaches zero for finite E, we may also write this as,

$$G = -kT \ln \int\int_0^\infty \Omega(V,E,N)\left\{\exp -\frac{1}{kT}(PV+E)\right\}\frac{P}{kT}\, dV dE. \tag{8.8.18}$$

We may extend the cases to include movable walls permeable to subsets of molecules. The last, and amusing, but unreal case that we consider movable walls permeable to all the molecules present, finally gets us into trouble. We sum W_K of (3) over all states of all energies, volumes, and number sets N to write

$$1 = \sum_N \int\int_0^\infty \Omega(V,E,N)\cdot\left\{\exp -\frac{1}{kT}[PV + E - N\cdot\mu]\right\}\frac{P}{kT}\, dV dE. \tag{8.8.19}$$

The sum of integrals on the right diverges, since for fixed P and μ there are related values of V,E,N such that $PV + E - N\cdot\mu = TS(V,E,N)$ are the actual values of TS for the equilibrium system. Since Ω is $\exp(S/k)$ we are summing something of order unity over all N to infinity.

The reason for the divergence is not hard to find. In using the differential expressions to derive (3) we had to assume V,E, and N negligible compared with the total values in the supersystem, namely, V_t, E_t and N_t. In all the other ensembles that were considered at least one of the extensive variables, either V, or at least one N_a, were held fixed. The probability, W_K, for states of very large values of the other extensive variables goes exponentially to zero. As long as V_t, E_t, and N_t are sufficiently large we are not doing violence to the equations by using the differential expressions, and the method becomes asymptotically exact as V_t, E_t, and N_t approach infinity in a fixed ratio, keeping T, P and μ constant. For the last ensemble no extensive variable was held fixed. We include in our systems those of any size for which the ratios of the extensive variables is the equilibrium ratio, and hence states of relatively high probabilities (the total number times the probability stays constant) even to infinite size. For fixed but large values of V_t, E_t and N_t, the expression (3) would be false for these states as V, E, N approached these values, as is obvious by the consideration that adding one more member to M_K would exceed the total size. The limiting process of going to infinity with V_t, E_t, N_t does not lead to a correct result.

8.9. Methodology of the partition functions

The scheme of ensembles, their partition functions, and the thermodynamic functions which they give is methodical and symmetrical. It is, however, clearest in a scheme which is not that usually used in thermodynamic nomenclature.

The usual thermodynamic functions and variables may be described as follows. The extensive variables are chosen as V, S, N. The internal energy, E, is a natural function of these variables, and it is this function, E, which is a minimum at equilibrium when V, S, N are held fixed. The partial derivatives, $-P$, T, and the chemical potential set μ,

$$(-P) = (\partial E/\partial V)_{S,N}, \tag{8.9.1}$$

$$T = (\partial E/\partial S)_{V,N}, \tag{8.9.2}$$

$$\mu_a = (\partial E/\partial N_a)_{V,S,N_b,\ldots}, \tag{8.9.3}$$

are then conjugate intensive variables to V, S, and N, respectively.

One then constructs functions,

$$H(-P,S,N) = E - (-P)V, \tag{8.9.4}$$

$$A(V,T,N) = E - TS, \tag{8.9.5}$$

$$G(-P,T,N) = E - (-P)V - TS, \tag{8.9.6}$$

$$\Theta(V,T,\boldsymbol{\mu}) = -PV = E - TS - \boldsymbol{\mu}\cdot\boldsymbol{N}. \tag{8.9.7}$$

Each of these functions is a minimum at equilibrium if its "natural" variables are held fixed, and the partial derivatives,

$$[\partial H/\partial(-P)]_{S,N} = -V,\ (\partial H/\partial S)_{P,N} = +T,\ (\partial H/\partial H_a)_{P,S,N_b\ldots} = +\mu_a, \tag{8.9.8}$$

$$(\partial A/\partial V)_{T,N} = +(-P),\ (\partial A/\partial T)_{V,N} = -S,\ (\partial A/\partial N_a)_{V,T,N_b,\ldots} = +\mu_a, \tag{8.9.9}$$

$$[\partial G/\partial(-P)]_{T,N} = -V, (\partial G/\partial T)_{P,N} = -S, (\partial G/\partial N_a)_{P,T,N_b,\ldots} = +\mu_a, \tag{8.9.10}$$

$$(\partial\Theta/\partial V)_{T,\mu} = +(-P),\ (\partial\Theta/\partial T)_{V,\mu} = -S,\ (\partial\Theta/\partial\mu_a)_{V,T,\mu_b,\ldots} = -N_a, \tag{8.9.11}$$

are all equal to the conjugate variable, with the plus sign for the intensive variable and the minus for the extensive.

Now another scheme is equally symmetrical and logical, but less convenient for the interpretation of directly measurable experimental quantities. Let V, E, N be the extensive variables, in which case the natural function $S(V,E,N)/k$ is now a *maximum* at equilibrium for constant V, E, N. The conjugate intensive variables are now

$$\phi = (kT)^{-1}P = k^{-1}(\partial S/\partial V)_{E,N}, \tag{8.9.12}$$

$$\beta = (kT)^{-1} = k^{-1}(\partial S/\partial E)_{V,N}, \tag{8.9.13}$$

$$\nu_a = -(kT)^{-1}\mu_a = k^{-1}(\partial S/\partial N_a)_{V,E,N_b}. \tag{8.9.14}$$

Following the scheme of eqs. (4) to (7) one may now construct natural functions for combinations of the extensive and intensive variables. The first, corresponding to H in (4), would be

$$(S/k) - (kT)^{-1}PV = \Phi(\phi,E,N), \tag{8.9.15}$$

$$(\partial\Phi/\partial\phi)_{E,N} = -V, \tag{8.9.15$'$}$$

$$(\partial\Phi/\partial E)_{\phi,N} = (kT)^{-1} = \beta, \tag{8.9.16}$$

$$(\partial\Phi/\partial N_a)_{\phi,E,N_b} = -(kT)^{-1}\mu_a = \nu_a, \tag{8.9.17}$$

which hardly seems to be a very valuable function. The others, on which $\beta = (kT)^{-1}$ replaces E, are less awkward. They are

$$(S/k) - (kT)^{-1}E = -A/kT = \Psi(V,\beta,N), \tag{8.9.18}$$

$$(S/k) - (kT)^{-1}(E + PV) = -G/kT = \Xi(\phi,\beta,N), \tag{8.9.19}$$

$$(S/k) - (kT)^{-1}[E + \boldsymbol{\mu}\cdot N] = -\phi V = PV/kT = \Theta^*(V,\beta,\boldsymbol{\nu}). \tag{8.9.20}$$

These functions (as well as Φ) are all maxima at equilibrium for changes at constant value of their natural variables. Their partial derivatives with respect to any of their variables are the conjugate variable, again with the negative sign for the extensive variable.

The ensembles of statistical mechanics are then entirely symmetrical and simple when described in terms of these variables and functions.

The value of each of these functions is given by the logarithm of the partition function for the ensemble of systems with fixed values of the natural variables. When one of the natural variables is intensive, rather than extensive, the partition function is obtained by summing (or integrating) over all values of the extensive variable after multiplication with a weighting factor. The weighting factor is always the exponential of the negative intensive times the extensive variable, namely

$$\exp(-\phi V) = \exp(-PV/kT) \text{ or } \exp(-\boldsymbol{\nu}\cdot\mathbf{N}) = \exp + (\boldsymbol{\mu}\cdot\mathbf{N}/kT).$$

We have not displayed the partition function for Φ, since it appears to have little practical value, but its formal construction is obvious enough.

The partition function, Ω, for the entropy, $S = k \ln \Omega$, is the simplest in appearance, mainly because the variables, V, E, N, have simple significance in terms of the mechanical coordinates and momenta. Those which are functions of T, $\boldsymbol{\mu}$ or P appear more complicated in form, but are actually often easier to compute than Ω.

Now, actually, although the formal description is neatest in terms of the functions, S, Ψ, Ξ, and Θ^* with paired variables (V,ϕ), (E,β), and $(\mathbf{N},\boldsymbol{\nu})$ it is more conventional to use the functions $-kT\Psi = A$, $-kT\,\Xi = G$, $-kT\Theta^* = -PV$, and the variable pairs $(V,-P)$, (S,T), and $(\mathbf{N},\boldsymbol{\mu})$. The functions A, G, $-PV$ are then $-kT$ times the logarithm of the partition function of the ensemble with the appropriate variables. These are the functions which are minima for variations at constant values of their natural variables.

In this hierarchy the equation $S = k \ln \Omega$ or $-TS = -kT \ln \Omega$ is unique, in that here the *variable* S appears as given in terms of the partition function computed for fixed value of the *natural function*, $E(V,S,\mathbf{N})$.

8.10. The general entropy expression, and the increase of entropy

For all the equilibrium ensembles the entropy, S, is given by

$$S = -k \sum_K W_K \ln W_K. \tag{8.10.1}$$

This follows directly from equation (8.8.3), that

$$-k \ln W_K = (PV - \mathbf{N}\cdot\boldsymbol{\mu} + E_K)/T. \tag{8.10.2}$$

The ensembles differ only in the limits on the quantum states $\boldsymbol{K}$ which are included. In the microcanonical ensemble only states $\boldsymbol{K}$ of fixed $V,\boldsymbol{N}$, and with $E \leqslant E_K \leqslant E + \Delta E$ are to be summed over. The extensive quantities V, N, E are those of the system, and

$$(PV + \boldsymbol{N}\cdot\boldsymbol{\mu} - E)/T = S. \qquad (8.10.3)$$

The sum in (1) is just $S\Sigma W_K = S$, since the sum over W_K is unity.

In the other ensembles E_K, V, or $\boldsymbol{N} = N_a, N_b, \ldots$, or any number of these quantities except one, may be different for different states $\boldsymbol{K}$. In any case the sum, such as

$$\sum_K E_K W_K = \bar{E}, \qquad (8.10.4)$$

gives the average value of the extensive variable for the system. The sum on the right of (1) is always $(-PV + \boldsymbol{N}\cdot\boldsymbol{\mu} - E)/T = S$, when the average values $\bar{V}$, $\bar{\boldsymbol{N}}$ or $\bar{E}$ are identified with the thermodynamic variables, V, $\boldsymbol{N}$, or E.

An ensemble of systems, not necessarily at equilibrium, would always be completely defined if W_K were specified for all states $\boldsymbol{K}$. For such an ensemble the average value of any physical property is determined (section 1.3). In general W_K would then be a function of time, a physical property would change with time, and we know empirically that it approaches the equilibrium value, and that the thermodynamic entropy will increase.

However, the initial *thermodynamic* entropy of the non-equilibrium ensemble can only be defined if one can describe, throughout the system, a *local* value to the intensive variables P, T, $\boldsymbol{\mu}$, which means that locally there must be equilibrium in the different degrees of freedom. We may, indeed, extend this concept to cases where, at one position, $\boldsymbol{R}$, two different degrees of freedom, for instance nuclear spin and lattice vibrations, have slow exchange of energy and differing temperatures; but thermodynamic entropy could not be defined if the distribution of energies within degrees of freedom which exchange energy extremely rapidly were far from that of equilibrium. In other words we must be able at least to conceive of adiabatic experiments measuring the entropy of the parts of the system, either by making the experiments more rapidly than the approach to equilibrium, or by inserting some sort of insulating walls that freeze the non-equilibrium configuration. In all such cases, since the entropy is the sum of the entropy of the parts (section 5) equation (1) will give a value consistent with the thermodynamically measurable entropy. We hence may use (1) as a general definition of entropy for all systems.

It would take us too far from the field of this volume, equilibrium statistical mechanics, to discuss fully the difficulties and answers to these difficulties with this definition. A few dogmatic remarks, however, appear to be in order. The quantity S defined by (1) can readily be shown for an ensemble of completely isolated systems to be stationary in time, that is *not* to increase. A common solution† is to define an average $\bar{W}_K$ obtained by averaging W_K over all quantum states $\boldsymbol{K}$ in the "neighborhood" of $\boldsymbol{K}$, and to define

$$\bar{S} = -k \sum_K \bar{W}_K \ln \bar{W}_K, \tag{8.10.5}$$

which now, with certain assumptions about the nature of the averaging process, can be shown to increase with time.

We prefer to use eq. (1) to define the entropy, but to emphasize that the change of W_K with time should not be computed for a hypothetical ideally isolated system, but for a more realistic model in which random time dependent fluctuations occur at the walls. If the walls are to correspond to a thermodynamically isolated system the fluctuations must be such that they do not influence the change with time of measurable macroscopic properties of the system. However, they will wipe out the long range multiple correlations between large numbers of molecules into which the negative entropy of the initial non-equilibrium macroscopic state flows. There is thus a real smoothing produced in the experimental systems by nature, and not introduced artificially by an unreal mathematical artifice.

In many or most experimental situations the rate of increase of S as defined by (1) using a realistically evaluated W_K would be numerically equal to the thermodynamically measured entropy. This is because the "negative entropy flow" into terms involving correlations between essentially all of the 10^{20} molecules of the system is rapid, and these correlations will certainly be destroyed by random wall events. However, insofar as true isolation and identity of members of an ensemble can be approached the S defined by (1) may indeed increase more slowly than the value of entropy inferred from the rate of change of the macroscopic parameters.

An example of this has been pointed out by John M. Blatt, to be given in the spin-echo experiment of Hahn. A simplified caricature

---

† See for instance R. C. Tolman, *The Principles of Statistical Mechanics,* Oxford University Press (1938).

of the experiment may be described as follows. Nuclear spins are lined up at low temperature in the z-direction so that the total magnetic moment is M_0. A weak homogeneous magnetic field in the x-direction imposed at $t = 0$, then causes the spin axes to precess in the y–z plane. Due to imperfections in homogeneity the precession rates are not identical, and after many complete revolutions the macroscopic magnetic moment, $M(t)$ not only rotates but decreases in magnitude. By a clever trick Hahn pulses with a magnetic field along the z-axis at time t_0 changing the angle, ϕ_i, that each spin i makes with the z-axis to $-\phi_i$, so that, continuing to precess at the same rate in the weak x-direction field, each spin returns to alignment on the z-axis at time $2t_0$, and the total moment to (nearly) its initial value, M_0. Since one infers a negative entropy increment proportional to the square of the macroscopic magnetic moment, $|M(t)|^2$, there appears to be an entropy decrease in the period t to $2t$, after which the normal increase again begins as the moment magnitude proceeds to the normal decrease.

Now first we must hasten to point out that careful attention to the rules of thermodynamic measurement removes any apparent conflict with the second law. A thermal measurement of S of the system of moment $M(t_0)$ at t_0 will be irreversible, since the system brought back on the supposedly reverse path to $M(t_0)$ will no longer have the property of increasing its moment to the initial value, M_0, under the x-directed magnetic field. A truly reversible thermal measurement of the entropy at t_0 can be made by "flipping" it to increase the moment to the value M_0 at $2t_0$, and then carrying out a reversible thermal measurement of the entropy. The measurement will now be reversible, since returning the system to the magnetic field with moment M_0 it will, after a time t_0 be at the initial state $M(t_0)$. The entropy will correspond to the normal entropy of magnetic moment M_0, which is much less than that inferred from the value $M(t_0)$. However, since actually the return is not perfect, $|M(2t_0)| < |M_0|$, there is a slow increase in entropy of the system, but much slower than that inferred from the decrease between $t = 0$ and t_0 from M_0 to $M(t_0)$.

The behavior has no mystery, but it does point out the danger of assuming an artificial smoothing of W_K by mathematical artifice, as implied in eq. (5). In the time between $t = 0$ and t_0 the slightly unequal precession rates of the different spins causes them to get out of phase, and a macroscopic measurement of the total moment $M(t)$ might lead us to infer that the rates were purely random, and should be "smoothed." Actually the rates are correlated with the field in homogeneities which stay constant in time, thus making it possible to recover, by the spin flip, the original uniformity of alignment.

The system is, however, an unusual one. Nuclear spins at close to absolute zero have notoriously small interactions, either with each other, or with the surrounding lattice vibrations. The experiment itself is also a rapid one, and the times, t_0, are far less than the relaxation times in, say, a normal macroscopic diffusion experiment at room temperature.

Blatt has compared the Hahn spin echo with the paradox proposed by Loschmidt (1872), namely that if, in an isolated system proceeding toward equilibrium, all momenta were exactly reversed, the system would run backward away from equilibrium. In order that all collisions should exactly reverse their course for a time t_0 requires that the correlations between all molecules at t_0 which were at the ends of all collision chains initiated by a single colliding pair at $t = 0$ should retain their correlation. These comprise most of the molecules within a sphere of radius ct_0, with c sound velocity in the medium. It seems inconceivable that at room temperature a normal system would reverse itself for more than a few milliseconds even were the momenta exactly reversed. The negative entropy stored in those correlations which have not been effectively smoothed by the actual time dependent random interactions at the walls would then correspond to the entropy change in the system during these few milliseconds. For a great class of cases this would be a negligible amount, and for these cases the entropy given by eq. (1) will correspond numerically to that inferred from the value of the local macroscopic state. For other systems a clever experimenter might always be able to show a method, as Hahn has done with the nuclear spins, of recovering the stored correlations, and experimentally establishing the lower entropy increase rate given by (1).

8.11. On logic and mathematical rigor

There is probably no other inclusive field of science in which it is more tempting to expect complete mathematical rigor from beginning to end than in equilibrium statistical mechanics. The axioms are the laws of mechanics, which, for molecular systems at least, can be put in concise mathematical form. The end product is an equally concise set of a very few (2 or 3) mathematically formulated laws, those of thermodynamics. It appears to this author, however, that a search for complete rigor in the usual mathematical form is illusory, and, when pursued too industriously, has more often led to obscurantism than to clarity.

The first difficulty that arises is that the conciseness and precision of mathematical formulation of the laws of thermodynamics are actually invalid for real systems of finite size; average values are not identical with the most probable, the probability of a spontaneous measurable decrease in entropy is infinitesimal but not zero, and phase changes are not singularities but merely finite changes in derivatives within an inobservably infinitesimal, but non-zero, range of variables. The demand that one treats only infinite systems obviates these difficulties, but leaves precious little applicability of thermodynamics to the real world. It also makes more awkward the extension to surface phenomena, or to edge effects in crystals, which are problems that, although not dealt with in this book, are fit subjects for statistical treatment.

The point of view adopted in this volume is to treat finite but large systems, and to show that the predictions of thermodynamics are, within experimental error, justified.

The nineteenth-century European scientists largely assumed that the object of a statistical theory would be to make proofs for the time average behavior of one isolated system. The ensemble concept of Gibbs is logically different. For many years a considerable mathematical effort was made to prove the two approaches to lead to identical results, an effort that culminated in proofs by von Neumann† and by Birkhoff‡. The time average behavior of a truly isolated system is a good classical concept, but hardly that which corresponds to actual measurements. A real macroscopic measurement disturbs the system, and does so far more than the minimum limit required by quantum mechanics. After one measurement this system starts afresh on a path in phase space having little or no relation to that which preceded the measurement. The ensemble of Gibbs appears to correspond much more realistically to experimental conditions than the time average behavior of one isolated system.

Similar difficulties plague the ergodic hypothesis. It has been proved§ that a classical hard sphere gas is ergodic, except for states of zero weight. The proof is a major mathematical accomplishment.

---

† J. von Neumann, *Proc. Nat. Acad. Sci. U.S.A.* **18**, 70 (1932).

‡ G. D. Birkhoff, *Proc. Nat. Acad. Sci. U.S.A.* **17**, 656 (1931).

§ See report of the Copenhagen Conference on Statistical Mechanics, July 1966, contribution by D. V. Shirkov, to be published by North Holland Press.

However, on the one hand the model of periodic boundary conditions is, just for this problem, unrealistically requiring a more difficult proof than that for a real system in which time dependent random fluctuations occur at the walls. On the other hand were this the only system to which statistical mechanical methods were valid the value to science in general would be very limited. A more realistic approach, and that adopted in this volume, would appear to be to ask, for each individual case, (A) what model must be used, such that if ergodicity within the states is assumed, the model does conform to the experimental system, and (B) what consequences an error in the assumption would have.

In the absence of extraordinarily detailed knowledge of the molecular mechanics in the system we can hardly expect an *a priori* answer to (A) in systems of any complexity. We do have a simple expedient to obviate the difficulty, namely to ask exactly what, and how many, macroscopic variables are necessary to define the state of the system. If the experimental systems are truly reproducible, then, as in the case of the CO crystal discussed in section 8.4, the best model to use is one which assumes ergodicity within the quantum states of the fixed macroscopic variables, whether or not transitions actually occur between all states.

The one exception, in which this procedure would be false, is one in which *all* experimentally prepared systems actually occupy only some distinct subset of fraction f (with $N^{-1}\ln f \sim$ order unity) of the quantum states assumed to be available. It would appear to be unlikely that this would happen without a reason that some clever theoretician would anticipate. It means, of course, that some hidden unsuspected macroscopic variable x exists, for which all known preparations of the systems lead to only a single value.

INDEX

References are to page numbers; bold-face numbers are usually the more important references, frequently the beginning of a several-page discussion